

Where the world stands on ozone

The Ozone Trends Panel which reported last week has confirmed that the expected downward drift in stratospheric ozone is under way, but not as yet catastrophically. The next three years will tell.

FOR most of the world, stratospheric ozone is understandably an abstraction. That there is ozone in the stratosphere at all is, for most people, simply hearsay; they have to take other people's word for it. And although the first measurements of stratospheric ozone go back more than half a century, for most of that time the measurements have been so unreliable that even those responsible have not known what to make of them. For the world at large, stratospheric ozone came into its own only when, in the early 1970s, the eventually successful lobby against a US plan to build a fleet of supersonic aircraft raised the scare (probably unfounded) that nitrogen oxides from burning fuel would reduce the average concentration of ozone.

The lobby succeeded not because of the threat to the ozone layer, but because the economic case against supersonic aircraft was overwhelming. But soon, more substantial — and permanent — mechanisms by which ozone, itself produced by solar radiation in the stratosphere, might be catalytically destroyed by the products of some of the artificial constituents of the atmosphere, the halogenated hydrocarbons in particular. To the world at large, it cannot but seem another occasion when unseen causes pose unconnected threats.

The report (see page 293) of the Ozone Trends Panel convened by the US National Aeronautics and Space Administration should go a long way to make the argument about the threat to natural ozone (and thus, by low-energy solar radiation, to people) at once more sober and less panicky. First, it is no longer possible to dispute that there have been some significant changes in the concentration of stratospheric ozone during the past two solar cycles (the present cycle has two or three years to go to its maximum). The theory that the changes observed are consequences of artificial chemicals such as halogenated hydrocarbons in the stratosphere is most convincingly borne out by rough correspondence between the predictions and observations at different latitudes. That is the bad news.

Good news

The good news, for the time being, is that the magnitude of the reduction of stratospheric ozone so far is comparable with natural variation between the trough and the peak of the solar sunspot cycle. So much is suggested by the expectation that, in the years before the present cycle peaks, the declining trend of stratospheric ozone in the Northern Hemisphere will probably be temporarily reversed. There could hardly be a more natural and convincing yardstick to tell what variations of transmitted low-energy ultraviolet radiation are physiologically tolerable. If the secular reduction of stratospheric ozone so far can be swamped by the effects of the sunspot cycle, it cannot so far have done much damage. The next solar sunspot cycle could, of course, be a different case.

The report of the Solar Trends Panel also vividly illustrates the difficulties confronting those who would more accurately monitor changes of ozone concentration in the stratosphere. A decade ago, it was generally hoped that ozone measurements would not thenceforth depend on ground-based measurements (mostly of the solar spectrum reaching ground level), but on instrumentally more sensitive and certainly more continuous

measurements from satellites. In the event, and a little ironically, the panel has been forced to the need to calibrate the satellite-borne instruments by the use of data from the much-despised Dobson network. This is not the first occasion when people trying to make sense of remote sensing measurements have been forced back onto old-fashioned ways, and will probably not be the last. That is why one of the most convincing of the secular trends now reported is the decrease of temperature (about 4°C) in the middle stratosphere at mid-latitudes.

The lesson to be learned, ideally by the time of the review conference of the Montreal Convention due in 1991, is that there is an urgent need for better instruments that can be calibrated more reliably. To be fair, public sources of research funds have been generous in their treatment of research projects in the field during the past decade, but it would now make sense to give special attention to schemes for improving on the sensitivity of the measurements now being made.

Ozone hole

Two puzzles remain, of which the most conspicuous is the famous Antarctic springtime ozone 'hole', first recognized three seasons ago and successively more marked in succeeding seasons. This dramatic phenomenon is a special if instructive case. Events have shown that the marked reduction of ozone in the low-lying spring stratosphere within the polar vortex out to a latitude of 60°S is indeed associated with high concentrations of the chemical ClO long suggested as the principal catalyst of ozone destruction, which goes a long way to support the notion that artificial chemicals are responsible. But, naturally, if there had to be a phenomenon of this kind, it could not be better placed than in the Antarctic. The second puzzle is that so little is so far known of the means by which the halogenated chemicals are removed from the stratosphere. While some gloomy forecasts assume that chlorinated hydrocarbons remain permanently in the stratosphere once they have found their way there, that cannot be literally the case. Is it even possible that the Antarctic hole may help to remove ozone from the stratosphere? While people improve their instruments, they might also pay some attention to this question, certain to become more pressing as chlorinated hydrocarbons continue to accumulate.

Meanwhile, it is fortunate that there is now a convention on the Protection of the Ozone Layer. The US Senate ratified the treaty last week. Last week's report should persuade reluctant governments (and their chemical manufacturers) to comply with the modest restrictions so far required of them — to freeze production of chlorinated hydrocarbons for the time being, and to reduce them by a half a some later stage. What has emerged, in the extremes represented by those who say that the damage done by the disappearance of the ozone layer would be counteracted if people in sunny climates carried sunshades and those who regard any sign of global anthropogenic change as an offence against nature, is a sober case for caution. And while skin cancer (among people) is often curable, the convention now on the books is an invaluable precedent for the much more contentious document that will be needed if and when it becomes necessary to contain carbon dioxide emissions. □



Stations in space

International partners in the US space station are signing up years before they should.

THE space station on whose construction the US National Aeronautics and Space Administration (NASA) is embarking will soon be fully provided, for except in one respect — a purpose. At the outset, the intended overseas partners (Canada, the European Space Agency or ESA and Japan) should have formally undertaken their obligations more than a year ago, but that hope was postponed by the row over what the US Department of Defense would require of the space station, while the need for agreement was diminished by the prolonged grounding of the shuttle. Only now are the documents being signed and sealed (see page 293). And while some grown men and women will now sleep easier in their beds, knowing that one hurdle lies behind them, others involved in the long negotiations, chiefly from the partner delegations, will sleep less easily while asking to what they have committed their governments.

The row about the interest of the Department of Defense, which dates from the end of 1986, has been a red herring. From the start, the Pentagon made no secret of its wish to follow the development of the space station (some parts of which may be in place by 1993) with the intention of using it for carrying out experiments if that should seem prudent. The pace of the development of the Strategic Defense Initiative (SDI) at the end of 1986 naturally prompted the Pentagon formally to ask NASA to keep its interests in mind while negotiating the terms on which overseas partners would participate. At least one of the potential partners saw this as an opportunity to withdraw with honour, on the grounds that it had no brief from its electors to sponsor SDI research. In the event, NASA diplomacy has plainly won the day, keeping its partners sweet while leaving the Pentagon's interest intact. Time will tell whether the doubters would have been better served if they had simply asked what the whole enterprise is for.

Nothing is more eloquent of the long planning delays preceeding innovations of technology in space, and of the way in which they necessarily embody ambitions belonging to an era past, than the space station as now conceived. Its roots lie in the period when it seemed as if the US space station had become just another means of transport, as if it were a kind of motor-car, and when NASA was casting around for other things to do. Reports of Soviet achievements with *Soyuz* spacecraft came thick and fast (as they still do) in 1984 and 1985, with the result that NASA naturally turned to emulation.

With the passage of time, two things have happened. First, ingenious people have devised ingenious uses to which a habitable space station might be put. The opportunities of microgravity (when thermal convection is suppressed) for fabricating pure crystals or separating materials cleanly by electrophoresis are so well canvassed that honest observers must often ask whether SDI does not have a better case. The plain truth is that the space station exists, if only as a concept now, because four and three years ago it seemed the natural next exploit. Unsurprisingly, it has been overtaken by events. Not only has there been the Challenger accident, but the United States has formally declared its aspirations towards even more ambitious goals in space — the inhabited base on the Moon and, sometime early next century, the trip to Mars. NASA may say that, providentially, the space station is just the means that a far-sighted person would have devised for getting to those distant places, but talk like that cuts no ice elsewhere.

The tragedy of the space station, as now conceived and voted for by participating governments elsewhere, is that it emphasizes the emptiness of fashionable interests in the exploration of space. It is not as if there is nothing else to do. When the Hubble Space Telescope is eventually launched (it should have been a going concern for at least a year), our knowledge and under-

standing of the Universe will be increased as much as it has been since Galileo's time: is that not a prize worth winning? Might it not be worthwhile, even at this poignant time when the shuttle is still on the ground, to build a second version in case the first should fail? Others will naturally have different but equally sustainable opinions of how the resources might be better spent on other ventures beyond the atmosphere.

More mundane considerations also obtrude. The cost of building the space station will be large, but not huge. The total cost will be roughly ten per cent of the annual federal deficit, but spread out over nearly a decade. Moreover, as NASA has always said, the money will not leave the ground, but will stay on the surface of the Earth, creating jobs and augmenting economic activity.

The flaw in that argument is that a country such as the United States, which has stumbled into the dubious role of the world's largest debtor, would ordinarily be compelled by economic necessity to persuade its creditors that it is a prudent manager of its own affairs. In the circumstances, may it not be particularly tactless that it should have set out to persuade its creditors to join in the folly? It is as if a man on the edge of bankruptcy tempted his bank-manager to the races for the afternoon. □

Living with civil war

This week's roster of assassination may, with luck, be enough to turn assassins' stomachs.

THE past few days have been shoddy times for the notion that the first element of freedom is that one's life will not wantonly be taken away. True, the Supreme Court of the Republic of South Africa has allowed that six people taken apparently at random from a murderous (and murdering) mob should have a stay of execution pending a retrial. But an Israeli conscript has been killed in Bethlehem, as have been more than a score of Palestinians in the West Bank during the past month, while we still have to learn with certainty what went on in Azerbaijan. Violence, as they used to say in the musical play *Oklahoma*, seems to be breaking out all over.

No more so than in Ireland, where the concept of death has been exquisitely refined (James Joyce's prose poem *The Dead* refers): if to die in a good cause is to be assured of a place in other peoples' memories, and perhaps even of immortality, if not literally then in the jungian sense (of the collective unconscious), may it not have been rational that the mourners at last Saturday's funeral in Belfast should have murdered the two soldiers who appear to have blundered into their way? What does it matter that two young people should have been, for the time being, denied the chance to chat up their friends?

The English, a subset of the British, do not understand the Irish notion of death, which is probably why a group of British soldiers considered that it made sense to kill rather than arrest three Irish people (one a woman) in Gibraltar two weeks ago, on suspicion (probably well founded) of planning to explode a bomb. Clinically, killing would have made the problem go away; in reality, it has ensured its persistence.

What happens now if the British government discovers there is no way out but the rule of law? Then there will be civil war in Ulster, not for the first time. Emollient devices, such as better housing or education, are no longer sufficiently immediate. Nor are schemes for pumping money into job-creation. Naturally, it would help if those people who vote in Britain and whose energies are spent battling *apartheid* in distant South Africa were to lend them to this more urgent cause. Almost certainly, that is asking too much. The discomfiting characteristic of domestic civil war is that it is uncomfortable; easier to fight somebody else's. Yet events are at such a pass that it may be permissible to hope that people will have been so sobered by what they have done that they acknowledge they have no choice but to find a better way to a different world. □

Politics and finance delay US space station

- Partnership problems still unresolved
- Price tag too high even for Japanese

Washington/Tokyo/Paris

THE European Space Agency (ESA) council last week approved the memorandum of understanding that defines European participation in the space station. As Canada has already come to an agreement, this makes Japan the only partner in the project yet to reach that stage, and could leave it out of discussions beginning later this month on specific hardware requirements.

Both Canada and ESA participation in the space station have been covered by extensions of earlier memoranda that had been due to expire at the end of last year. But the National Aeronautics and Space Administration (NASA) extended the agreements pending formal approval.

In fact, ESA will not sign the new memorandum until all its member governments sign specific intergovernmental agreements with the United States. Although these agreements have also been worked out in principle, ratification will take some time.

For the Japanese, a problem has arisen between the Science and Technology Agency (STA) and the Foreign Ministry over which agency should sign the memorandum. Masahiro Kawasaki, director of STA's research and development bureau, says there is no dispute, and that "coordination discussions" should resolve the issue by early April. But NASA is not sufficiently convinced that the problems are minor, and for the time being Japan is excluded from participating in any space station meetings. This could mean that when Japan does join in, it will have to accept the engineering configurations the other partners decide upon.

Even in the case of Japan, the fundamental problems that had held up international participation — most notably the Pentagon's wish to keep open the option of using the space station for military purposes — have now been resolved, but financial problems remain. The Reagan administration's 1989 budget seeks \$1,000 million for the space station, and an overall increase in the NASA budget of \$2,500 million. But the House of Representatives Budget Committee has already indicated that NASA's increase is likely to be only half that large, and some of that reduction will surely come out of the space station's allocation.

Budget problems are also threatening Canada's participation in the space station. Canada's financial contribution to

the project had originally been set at \$C800 million. But fluctuating currency exchange rates and refinements in hardware specifications have boosted that figure to approximately \$C1,200 million. An Ottawa decision to place a strict cap on spending has also meant that difficult choices must be made on how to provide the extra money. The Canadian government has refused to rule out the possibility of withdrawing from the project entirely.

Canada faced particular problems in working out its memorandum of understanding for participation in the space station. ESA's man-tended free flyer, attached pressurized laboratory and polar orbiting platform and Japan's experiment modules with pressurized and unpressurized segments, can all be isolated from the rest of the station, should a proposed military use not be acceptable.

But the Canadian Mobile Servicing Unit is part of the station's infrastructure. Establishing a dispute resolution system was a prime concern for Canada, and in practical terms Canada will find it difficult to enforce its point of view once the station is in orbit.

The current reorganization of governmental science policymaking agencies has also complicated Canada's space posture. At the moment, Canada has no space agency. Instead, space activities are distributed among several ministries, including Communications, Engineering and Mines, the Ministry of State for Science and Technology and the National Research Council. There has been a long-standing plan to form a space agency, and a transition team led by Arthur Collin was formed to that end. But after more than a year there is still no clear plan for which ministry should be the permanent home of the space agency, and Collin has resigned. There is even a debate over the location of the agency, with a regional dispute started by Montreal to wrest it from Ottawa.

Further to confuse matters, Canada is merging its Ministry of State for Science and Technology into a larger department that will be headed by Industry Minister Robert de Cotret. Two weeks ago, de Cotret indicated there might not be a space agency if Canada does not participate in the space station. But last week, his junior colleague, Science Minister Frank Oberle, seemed to be saying that the space agency would be set up anyway.

Joseph Palca, David Swinbanks
& Peter Coles

CFCs cause part of global ozone decline

Washington

ATMOSPHERIC ozone levels at middle and low latitudes decreased by 2.5 per cent between 1978 and 1985, and some part of the depletion has been caused by chlorofluorocarbons (CFCs), according to the Ozone Trends Panel, a committee set up to re-examine existing data from satellites and ground stations. These findings, announced last week, are less alarming than earlier estimates of a general decrease as large as 1 per cent per year, but nevertheless cannot be explained solely as the result of changing solar activity.

The Ozone Trends Panel (OTP), an international group of atmospheric scientists, was set up in October 1986 by the National Aeronautics and Space Administration (NASA) and other agencies in response to claims made before Congress by Donald Heath, of the NASA/Goddard Space Flight Center, that ozone levels were decreasing globally at an inexplicably high rate. Heath derived his results from satellite measurements only, the difficulty being that a known deterioration in the instruments had to be accounted for to obtain the true ozone trend. By contrast, OTP compared the satellite data with measurements from several ground stations to deduce a more reliable calibration, and arrived at an ozone loss rate of 2.5 per cent from 1978 to 1985.

According to OTP, anything from 0.7 per cent to 2 per cent of this total could be the result of diminishing ultraviolet radiation from the Sun, which reached the minimum of its eleven-year sunspot cycle in 1985. Robert Watson, project manager at NASA, is now certain that trace amounts of CFCs are responsible for the remainder of the decrease. As the Sun's activity increases towards Solar Maximum in 1991, ozone levels may hold more or less constant, but then the decline will resume.

A more worrying part of OTP's report is that at high northern latitudes, between 50° and 60°, ozone concentration has fallen a little faster, to the extent of 6 per cent since 1970. This is too much to be explained by gas-phase CFC reactions alone, and may imply that a smaller version of the Antarctic ozone hole, which is instigated by reactions on ice particles, is appearing around the North Pole. This new finding, says Watson, adds impetus to investigation of the atmospheric chemistry of the Arctic (*Nature* 331, 201; 1988) because it is important to discover if the mechanism causing the northern depletion might be of global significance.

David Lindley

Social sciences undervalued?

Washington

OPPORTUNITIES for scientific advance in the behavioural and social sciences are being neglected, according to a new report from the US National Research Council (NRC), the operating arm of the National Academy of Sciences.

Funding for these subjects has gone up and down like a roller coaster following a surge in the 1960s, but the downs have been bigger than the ups. Between 1972 and 1987, federal support fell 25 per cent in constant dollars, while support for other areas of science rose 36 per cent.

Behavioural and social sciences encompass a huge range of research areas, from brain structure and neurotransmitters, through emotion and motivation, to the structure of human institutions, including organizations that decide science policy. The report maps the whole territory, concentrating on the areas where a push could produce rapid advance (see right).

The needed push requires human, technical and financial resources. But the resources are not there and the report complains of a "persisting view" that behavioural and social sciences research "is a virtually equipment-free enterprise" for which little more is needed than a "slightly larger office allocation". Even the tongue-twisting title of the National Science Foundation's Directorate for Biological, Behavioral and Social Sciences is thought to suggest a second-class status.

According to the report, behavioural and social scientists have almost as much need of computers as do physicists. At the behavioural end of the spectrum they make it possible to construct images of the brain from scanners; at the social end they permit the manipulation of large-scale arrays of social and economic data.

The report recommends an extra \$51 million a year for computers, neuroimaging devices and laboratory equipment. \$40 million extra per year is earmarked for the construction of large-scale databases, and \$10 million for access to government and corporation data collections.

On top of that comes a request for \$44 million to improve training at the pre- and postdoctoral level as well as at workshops; and \$70 million for individual investigator grants. A perceived need for far more interdisciplinary studies generates a request for \$25 million for new centres.

The whole bill comes to an increase of \$240 million, a 13 per cent rise on the current annual budget of \$778 million.

Alun Anderson

The Behavioral and Social Sciences: Achievements and Opportunities, is published by the National Academy Press, Washington, DC, 1988.

Yes, social sciences really do matter

Washington

THE NRC's report of US social sciences, even though at a coarse resolution of 23 headings, shows too many hundreds of research opportunities to try to list (see left). But with more money for parts of the behavioural and social sciences, we might know how to make the world a better place, or at least know who to blame.

People, it seems, have been doing unpleasant things to one another for a long time: the oldest archaeological evidence of organized warfare comes from a 12,000-year-old cemetery in the Nile valley (see 'The evolution of human society').

More accessible to immediate analysis is criminal behaviour in the United States. Half of all urban US males can be expected to be arrested for a non-traffic offence sometime during their lives (see 'Crime and violence'). To find out why, social scientists are supplementing police reports, the traditional source of data on crime, with confidential reports from criminals about their own activities.

Both data sources confirm that a small

number of 'career' criminals are responsible for a disproportionate percentage of crimes. While teenagers commit most crimes of all, the majority of them stop offending after adolescence. Career criminals continue through their 30s or sometimes even 40s.

What determines when people abandon crime? And how are criminal careers affected by punishment? Is it worth imprisoning some young criminals for long periods when their careers may soon come to a natural end anyway? Finding the answers to these questions requires large-scale surveys with overlapping cohorts of subjects. And that is expensive.

Some would like to blame everything on modernization and the concomitant decline of the extended family: analysis of historical archives is turning this view upside down (see 'Modernization: family and religion'). It seems that the nuclear family emerged at the closing of the middle ages and provided the independent people who originated the industrial revolution, rather than being its result.

Alun Anderson

International SSC funding causes scepticism in Congress

Washington

The Department of Energy (DoE) — anxious to find ways to persuade a financially tight-fisted Congress to commit to the expensive Superconducting Super Collider (SSC) — has promised that governments all over the world are enthusiastic about the project, and might contribute as much as half the estimated \$4,400 million cost of building it. But a congressional committee last week said that DoE's promises must be backed up by firm commitments, not good intentions.

James Decker, acting director of energy research, appeared before the international scientific cooperation subcommittee of the House of Representatives Committee on Science, Space and Technology to describe the interest other countries have in SSC.

In January, Decker and George Bradley, principal deputy assistant secretary, between them have visited Italy, France, the United Kingdom, West Germany, Switzerland, Canada and Japan, as well as CERN (the European laboratory for particle physics). Decker described these visits as extremely cordial, saying there was great enthusiasm for SSC, especially among physicists.

But under questioning from committee members, Decker was unable to describe any firm commitments to SSC. In budget documents delivered to Congress in February, DoE indicated that it expected foreign contributions, both cash and "in-

kind", to amount to \$1,800 million, but Decker was unable to offer much explanation about how the figure was reached. Part of the \$1,800 million is expected to come from financial packages put together by the state that wins the competition to be the home for SSC.

A classic Catch-22 predicament is in danger of developing with respect to international participation in SSC. DoE argues that a strong financial commitment from Congress — in the form of a hefty appropriation — will convince other countries that the United States is serious about the project. But Congress, wary of empty promises, has given notice that it wants a substantive show of support before committing itself. In testimony to Congress two weeks ago, Energy Secretary James Herrington said that foreign participation could amount to as much as 50 per cent of the cost, but Decker told the committee he thought 40 per cent was a more reasonable high-end figure.

For their part, the committee members seemed unconvinced by Decker's rosy evaluation of the situation. Representative Robert Roe (Democrat, New Jersey) said SSC had entered the "real world", and its supporters would have to provide realistic budget estimates if the plan was to be sold to the entire Congress. Roe also warned that foreign scientists might be prohibited from using SSC if their governments did not contribute to the cost of its construction.

Joseph Palca

Chemical vapour deposition advances superconductors

Tokyo

JAPANESE researchers are developing a technique that may open the way to mass production of wires, coils and complex-shaped thin films of high-temperature superconductors. A group at Tohoku University, in collaboration with Riken Corporation, a piston-ring manufacturer, has used chemical vapour deposition to coat a variety of substrates with thin films of yttrium-barium-copper oxide.

Chemical vapour deposition, widely used in the semiconductor industry, allows thin films of even thickness to be deposited over the micron-sized steps and channels of the circuit substrates, and large numbers of substrates can be coated at one time. Several semiconductor manufacturers around the world have reportedly tried the technique with high- T_c superconductors, but could not find a suitable barium compound to vaporize. Until now, physical vapour deposition techniques, such as sputtering and electron beam evaporation, have been used to prepare high- T_c thin films.

In the new method, evaporated 'source materials' (of undisclosed composition) containing yttrium, barium and copper are heated separately to different temperatures, evaporated and carried by argon gas into a heated low-pressure reaction chamber (700–900 °C, 10 torr) containing the substrate to be coated.

Oxygen is introduced separately, chemical reactions occur and a thin film of yttrium-barium-copper oxide is deposited on the substrate at a rate of about 1 micron per hour. Annealing is not required but oxygen treatment after deposition improves film quality.

The Tohoku team has coated a variety of substrates using the new technique, including a coil of chrome-nickel alloy. Unlike physical vapour deposition, for which particular substrates such as strontium titanate or magnesium oxide must be used to produce films with orientated crystals, the chemical vapour deposition method allows crystal size and orientation to be adjusted by altering the conditions in the reaction vessel.

According to Professor Toshio Hirai, highly orientated polycrystalline films have now been prepared and he foresees use of the technique in the manufacture of current-carrying wires and coils and in the development of electronic devices.

The technique is still in its infancy. Critical currents have not yet been measured and the critical temperature T_c at which electrical resistance vanishes is less than in thin films prepared by physical vapour deposition. But on 3 March, the team announced a T_c of 43 K and by last Thursday they were up to 72 K for a film on a magnesium oxide substrate.

David Swinbanks

CFC protocol agreed

Washington

THE US Senate last week voted 83–0 to approve ratification of an international agreement that was signed last September in Montreal by 31 countries, and is designed to limit global release of chlorofluorocarbons (CFCs).

The protocol now goes to the White House where President Reagan is expected to sign it. Mexico is the only other country to have approved the treaty, but the United States is the first major user and producer of CFCs to approve it.

The protocol takes effect on 1 January 1989 provided that 11 countries accounting for two-thirds of the global CFC consumption have approved it. Otherwise, the protocol becomes effective 30 days after that criterion is met.

J.P.

HIV-2 detected in UK

London

As was inevitable, a case of infection with the second type of human immunodeficiency virus (HIV-2), which predominates in West Africa, has been detected in Britain. The evidence has emerged from a small-scale programme of anonymous testing at Middlesex Hospital Medical School of blood samples taken in a sexually transmitted disease clinic in London. Cases of HIV-2 infection in other European countries, notably France and Sweden, and in the United States (see *Nature* 331, 381; 1988) have so far been confined to people with West African connections. Although that is almost certainly also true of the infected individual now detected in London, the anonymity of the testing prevents confirmation. Blood-bank screening for HIV-2, as well as HIV-1, may have to follow.

P.N.

Newborns to be tested

Washington

THE Centers for Disease Control (CDC) in Atlanta announced last week a programme for anonymous blood tests of newborns for the presence of antibodies to HIV (human immunodeficiency virus).

The tests will be conducted in 30 cities in the United States, chosen on the basis of a high reported prevalence of HIV, high prevalence of other sexually transmitted diseases and geographical distribution. The testing programme is expected to be ready by the end of May.

The newborn testing is just one of a series of surveys being conducted by CDC to obtain a better picture of the prevalence of HIV in the general population. Other surveys will include voluntary testing of patients visiting sexually transmitted disease clinics, as well as tests on adult hospital patients. The hospital tests will involve orthopaedic patients, who as a group are more representative of the general population than other in-patients.

J.P.

AIDS and public opinion in France

Paris

TWO surveys carried out recently by researchers at INSERM, the French National Institute of Health and Medical Research, show how information campaigns about AIDS are affecting public opinion and behaviour. The first, carried out last June had a nationwide sample of 1,500 respondents; the second, carried out in December 1987, sampled 900 people in the Paris area. Although the two surveys are not strictly comparable, they show a change in attitudes to AIDS over the six months between them.

According to both surveys, AIDS is considered a major health risk by the French public — less feared than cancer, but more feared than heart disease. But the rate of false beliefs about the disease is high. Just over 52%, for example, believed they could contract the AIDS virus by giving blood and nearly 37% thought the virus could be transmitted through saliva.

Although less educated respondents felt they could contract the disease through simple bodily contact, even those with a *baccalaureat* (secondary education

diploma) felt that the human immunodeficiency virus (HIV) could be transmitted by mosquitoes, on dental instruments or by sharing a hospital ward with a patient suffering from AIDS.

Attitudes towards compulsory screening for the presence of HIV had changed between the two surveys. In June, 73% of the sample were in favour of systematic screening of the whole population, while in December the proportion had dropped to 37.8%. However, the researchers point out that many of the later sample were in favour of compulsory screening of certain sectors of the population. Over 78% felt that pregnant women should be screened.

The AIDS information campaign initiated by the health ministry has concentrated on short television films and films aimed at secondary school children. Leaflets have also been sent out to general practitioners. But, compared to Britain, where there are regular poster campaigns, magazine advertisements and films about the disease, it is possible to be in Paris for weeks without seeing any evidence of a health campaign.

Peter Coles

First Indian remote-sensing satellite launched by Soviets

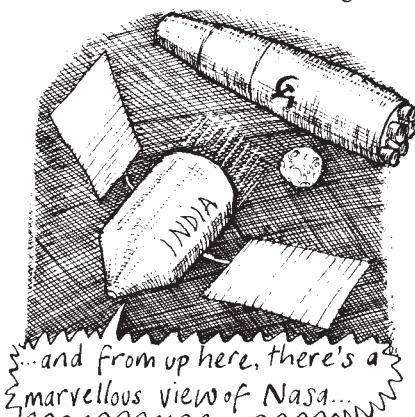
New Delhi

INDIA'S first operational remote-sensing satellite IRS-1 has gone into orbit after a successful launch from the Soviet cosmodrome at Baikanur on 17 March, an event Prime Minister Rajiv Gandhi described as "a major milestone in our space programme". With this, India has become the first developing country to have its own space-based system for natural resource management.

For the Soviet Union, the launch of the 980-kg Indian spacecraft was the first since its space business went commercial. The Indian Space Research Organisation (ISRO) will pay \$6.5 million for the launch, 25 per cent less than the price for similar services provided by the United States or the European Space Agency.

The \$50-million ISRO-built spacecraft

carries three French-supplied linear-imaging self-scanning cameras. From a height of 904 km, they can take pictures of 148-km wide scans in four spectral bands (0.4–0.9 micron) with a spatial resolution of 36 and 72. The polar Sun-synchronous orbit enables the satellite to revisit a fixed place on Earth once every 22 days. The IRS-1 data will be received at ISRO's ground station near Hyderabad, set up in 1979 to receive data from the US Landsat and augmented last year for reception from the French satellite Spot. Some 300 users, mostly in government agencies, have been identified for utilizing IRS-1



Indian scientists too dependent on West

New Delhi

Too much dependence on the West is one reason why Indian science is not able to tackle local problems, according to A.S. Paintal, president of the Indian National Science Academy (INSA). The academy has appealed to scientists to "think independently for ourselves" instead of working for honours from western nations which, says Paintal, maintains double standards — one for themselves, and another for Indians.

India has few institutions which can carry out research and development work in areas that should be receiving top priority. For example, Paintal says, nobody has thought of finding ways to reduce the temperature of workplaces to 28°C, which alone would improve efficiency. No technologies have been developed to tap India's huge hydroelectric resources and applied work pertinent to national problems has been neglected.

One consequence of dependence on the West is that Indian scientists and their work are evaluated by western yardsticks, says Paintal. "We think an Indian scientist is good only if some western organization thinks he is good." Paintal says that this attitude is being exploited by western organizations by giving honours to those Indian scientists suited to their own needs.

Paintal also warns Indian scientists to be wary of being used as "covers for activities of the developed world". "Indian scientists must learn to think independently and should no longer be allowed to slog for some honour or favour of dubious merit and uncertain use to the country."

K.S. Jayaraman

data products.

According to ISRO, IRS-1 is the first in the series of remote-sensing satellites it will put in orbit to form the space segment of the newly established national natural resources management system. The second IRS will also be launched from abroad but Indian's own rocket, PSLV, will launch IRS-3.

Operating the IRS system will cost India about \$80 million a year, a worthwhile investment in the context of increasing privatization of remote-sensing technology and possible disruption of services provided by foreign satellites. India is bitter about the rising cost of remotely sensed data products and considers the annual access fee of \$600,000 to Landsat as excessive. From ISRO's point of view, the launch of IRS-1 has brought India a step closer to its goal of self-reliance in a vital sector of the space industry that has so far been the preserve of the West.

IRS-1 cannot match Landsat or Spot in picture quality, but future satellites should carry higher-resolution cameras and a 'microwave eye' to see through clouds. ISRO believes India will capture a share of the remote-sensing market.

Space imaging has so far been used to map forests and identify wastelands, and the huge investment in remote sensing has yet to pay dividends for farmers. ISRO will not be ready before 1990 to apply this technique to more important tasks.

K.S. Jayaraman

Cash windfall for UK pioneers of MRI

London

Two British universities are this week grappling with the pleasing problem of how to carve up their share of more than £3 million in royalties arising from pioneering work on nuclear magnetic resonance imaging (MRI) over the past 15 years. The payments have been made to the universities of Nottingham and Aberdeen and to various individual scientists by the British Technology Group (BTG), a state-owned management company set up to help bring innovative ideas to the market-place. Money has been given to seven researchers with specific agreements with BTG. Precise sums are being kept secret.

Many of the basic techniques now used for diagnostic imaging were developed at the two universities between 1974 and 1980, with backing from the Medical Research Council and the universities themselves. The key inventions were patented by the National Research Development Corporation, now part of BTG. The main breakthrough on royalty payments was achieved in December 1986 when BTG accepted a substantial out-of-court settlement from the US healthcare company Johnson & Johnson for infringement of patent. Six months later, after

nearly two years of talking, BTG concluded an agreement with General Electric, of the United States, which included a down-payment of several million dollars — the source of the money BTG has just paid out to Aberdeen and Nottingham.

BTG has a portfolio of 16 separate inventions on MRI. The main patent, developed by Peter Mansfield's team at Nottingham, is for the 'slice-selection' technique, which allows selective excitation and imaging of a specific cross-section of sample. More money is expected to come BTG's way as MRI technology continues to evolve with the consequent expansion of the market. At present, each MRI system costs around £1 million and more than 600 are installed worldwide. The MRI market from next year is expected to be worth between £300 million and £600 million annually.

Within the two universities, decisions on precisely how to spend the money have yet to be taken. Aberdeen's options are narrowed by its policy of splitting royalty cash three ways, with equal portions going to the department involved, individuals (weighted according to their contribution — "which can be difficult" concedes a university official) and the university's central funds.

Simon Hadlington

Critics doubt five-year-old SDI programme will ever mature

Washington

A NEW report on the first five years of US government funding for the Strategic Defense Initiative (SDI) concludes that industry is still not sufficiently committed to the programme to give it "unstoppable momentum".

A companion report predicts that arguments over whether the SDI systems now under development can be tested without breaching the Anti-Ballistic Missile (ABM) treaty will continue for the remainder of the Reagan presidency. A combination of these factors should leave the next president with room to manoeuvre in deciding whether to deploy the beginnings of a missile defence system.

But President Reagan, speaking last week on the fifth anniversary of his speech proposing SDI, reiterated his support for

among a large number of different corporations. Many of them expect funding for other defence projects to fall in the 1990s as the military build-up ordered by President Reagan passes its peak. That may mean corporations will push hard for deployment of SDI so that they can fill their order books in the next decade. So far, however, industry has been less than enthusiastic about SDI.

SDI contracts have been awarded solely for research and have not generated big profits. These will come only if systems are ordered for deployment. But there are enormous uncertainties: several big projects have been delayed as they failed to live up to technical expectations and one, the McDonnell-Douglas project for a neutral particle beam weapon, has been cancelled outright.

Further uncertainty is added by continuing congressional opposition to tests of missile defences. The administration's 'broad' interpretation of the ABM treaty,

signed by the Soviet Union and the United States in 1972, would have permitted testing of SDI systems based on novel physical principles. But that interpretation of the treaty's language has not won acceptance in Congress. Now the administration is turning to a 'permissive' interpretation of the treaty that circumvents the ban on testing of components of a missile defence system by arguing that a component is not a component unless it could be part of a finished missile defence system.

John Pike, FAS associate director for space policy, caricatures the administration view as saying that "a car is not a car without a racing stripe". Challenges to the administration view are likely to come this year and early next year when tests fall due for the Airborne Optical System, a modified Boeing 767 carrying an infrared telescope for tracking and identifying re-entry vehicles while they are still above the atmosphere. The administration will argue that because the aeroplane can stay aloft only for ten hours, and because the infrared sensor chip it carries will not have a full array of elements, the system cannot be regarded as an effective 'component' of an ABM system.

Alun Anderson



the project. He criticized Congress for restricting SDI spending, arguing that the Soviet Union is actively pursuing strategic defences of its own. Although many physicists maintain that Pentagon estimates of the technical obstacles to SDI are optimistic, Reagan said that "if anything, we overestimated the technical challenge" of SDI.

The two critical reports come from the Federation of American Scientists (FAS), a non-profit organization representing some 5,000 physical and social scientists and engineers. The first shows that SDI is more a Californian project than a US one: \$6,600 million of contract money — 45 per cent of the total — is being spent there. Adding in New Mexico and Massachusetts accounts for two-thirds of all spending. Unlike the B1 bomber programme, which proved impossible to kill off, the SDI programme has not succeeded in awarding large contracts in every congressional district.

But SDI funds are very widely spread

CNRS outlines its biology plans to 1990

Paris

As French political parties preparing for the May presidential elections play 'double or quits' with electoral promises, the Centre National de la Recherche Scientifique — (CNRS) — France's largest state research organization — has been taking a more sedate look at its priorities for the life sciences in the 1990s.

The life sciences division is the largest within CNRS, its 250 laboratories and 1,000 researchers swallowing half of the annual budget. It has been nine years since François Gros, François Jacob and Pierre Royer set out the last perspective for French life sciences in their report to the president of the republic. In the meantime, the face of biology has changed dramatically and the CNRS administration has now promised to revise the new programme of priorities every year.

As an overview of contemporary trends in biology, the report contains few surprises. But the authors do pinpoint particular areas of strength in French biology and (with un-Gallic candour) some weaknesses. The four major 'social and economic stakes' picked out for the 1990s are painted with a very broad brush — health, agriculture, biotechnology and the environment. These areas, where biology can be seen to be relevant to the quality of life in France, are presented before the eight basic research themes.

France's role in isolating and describing the human immunodeficiency virus (HIV) has given a boost to immunology research.

Under the rubric of 'health', the report looks forward to progress in the study of oncogenes and the use of *in vitro* fertilization to study mechanisms in human reproduction and the diagnosis of genetic defects at the zygotic stage. In the fields of virology and vaccine development, the authors feel that the "small number of centres for sorting synthetic molecules" could be an obstacle to the study of antiviral molecules.

Apart from work to sequence the genomes of bacteria such as *Escherichia coli* and *Bacillus subtilis*, or of yeasts such as *Saccharomyces cerevisiae*, the report suggests that French laboratories could get a "ten-year lead" on US scientists by developing their efforts to sequence the genomes of "pseudo-unicellular" protista, as opposed to eukaryotes. But within enzymology, France has so far failed to exploit some of the latest techniques in X-ray crystallography, the report says.

The report is initially aimed at the CNRS administration committee and, does not address specific questions of finance. But, as a CNRS communiqué points out, one of the most daunting obstacles to future developments will be the high cost of major research programmes. Having been funded for decades with "artisan" budgets, major international programmes will require an investment "on a par with those for space or for research instruments in physics", says the pamphlet accompanying the report.

Peter Coles

A. Berkaloif, R. Naquet & J. Demaille *Biologie* 1990 — enjeux et problématiques. CNRS, October 1987.

Soviet reform of education

London

THE Soviet Union is to amalgamate into a new State Committee for Public Education the three All-Union bodies with responsibility for education — the ministries for education and for higher and specialized education and the State Committee for Public Education. The retiring Higher Education Minister, G.A. Yagodin, will head the new state committee.

The objective, apart from the avoidance of job duplication and bureaucratic barriers as part of *perestroika*, appears to be closer coordination between secondary and tertiary education. Reform along similar lines in the Estonian republic was announced only a few months ago.

Concern has been growing among Soviet educationists about the gap between the educational attainments of school-leavers and the minimum requirements of universities and colleges in the tertiary sector.

Part of the trouble seems to have been the natural desire of school examiners to qualify as many school-leavers as possible, but the lack of a nationwide clearing-house for students has meant that less well-qualified students have secured places in higher education at the expense of others by limiting their ambitions. Meanwhile, universities and colleges have tended to recruit less able students to maintain their numbers.

Closer links between schools and universities are a central part of the new trend in higher education, notably in the proposal that senior school students should have access to facilities in universities. In its new form, the link between school and higher education has practical rather than philosophical goals. It seems to be acknowledged that Soviet schools are poorly equipped for modern education (in computing, for example) as well as for their decreed role of work-education.

Although universities are hardly in better shape, the Central Committee seems to have calculated that the new arrangements under which universities will equip themselves by forming links with industrial enterprises will allow them to discharge their new responsibilities towards the schools.

The reorganization now under way nevertheless leaves many questions unanswered. Allowing for parallel ministries at the republic level, there will remain more than 70 ministries and other bodies with responsibility for education.

No central body has yet come to grips with the question of whether work-orientated ('polytechnic') education is the best preparation for all students embarking on tertiary education.

Vera Rich

University reorganization stirs controversy in Germany

Munich/Köln/Bonn

Conflict is brewing across North Rhine-Westphalia over a controversial plan to reorientate the universities towards high technology. Students and professors alike have been on the streets protesting against cuts in natural sciences, humanities and education courses. As tempers have risen, the ambitious education policy of the governing Social Democrats has increasingly been called into question.

The northern German *Land* (state) of North Rhine-Westphalia (NRW) is threatened by financial crisis. Huge budget deficits loom over the Social Democratic (SPD) government, as the region's steel and coal industries are rationalized.

The 'university structure plan 2001' was introduced last year with the aim

of increasing competitiveness with the rapidly advancing south German *Länder* by emphasizing university courses in electronics, computer science and other technical fields. At the same time, cuts were proposed in fields where there are already surpluses of graduates.

The ministry has been forced to retreat in areas where student and faculty protests have been intense. For example, NRW Science Minister Anke Brunn announced in January that the physics faculty in Bonn would be "pushed" while physics in nearby Köln would "not necessarily be eliminated". This brought the Köln physicists and their students onto the streets.

Responses to the plan have depended on individual institutions' histories. The four oldest universities — Köln, Bonn and Münster and the Technische Hochschule at Aachen — stand to lose the most. The newest institutions — the Gesamthochschulen in Duisburg, Essen, Paderborn, Siegen and Wuppertal — will almost certainly be protected. The reasons trace back to the ambitious expansion undertaken by the former science minister — and now NRW minister-president — Johannes Rau (SPD) in the 1970s. The philosophy was to "regionalize" NRW's universities by opening five institutions in new places, in addition to the four universities founded in the 1960s (in Bochum, Düsseldorf, Dortmund and Bielefeld).

The five newcomers have established themselves in useful niches, for example in microelectronics in Duisburg. They have brought economic benefits to previously underdeveloped areas, as in Siegen. The financial costs have been high but the political costs of closing even one of the universities would be even higher. It would be tantamount to admitting that the regionalization policy was a failure.

The greatest threat to the viability of NRW's universities is the expected reduction in the number of students by the mid-1990s, currently estimated at 30 per cent. In order to avoid the humiliation of having to close a Gesamthochschule, the Science Ministry is making every effort to make them more attractive by its restructuring plan.

The Ministry has promised not to make any outright cuts until after the next elections in 1990. But what will happen then? Unhappy at the ministry's autonomy in the restructuring process, a group of researchers have appealed to the Wissenschaftsrat, West Germany's federal science advisory council, for an outside appraisal. The council will decide this summer at the earliest if it wants to make recommendations. It seems likely that it will.

Steven Dickman

Natal court casts out government ruling

Oxford

THE Natal Supreme Court on 14 March declared the subsidy-linked disciplinary conditions imposed on the University of Natal to be of no force and effect. The conditions were imposed last October by the Minister of Education and Culture in the South African House of Assembly, Piet Clase (see *Nature* 330, 4; 1987). Mr Justice Page ordered Minister of National Education F.W. de Klerk, to pay the costs.

The Natal case had earlier been adjourned pending the release of the full judgement of the Cape Supreme Court in a similar successful application by the universities of Cape Town and the Western Cape (see *Nature* 331, 651; 1988). Counsel for the universities had attacked the validity of the conditions on three grounds: (1) their imposition exceeded legal powers; (2) they are so vague that they do not convey what the universities are required to do to avoid non-compliance; and (3) they involve unreasonably oppressive or gratuitous interference with the universities' rights. Mr Justice C.T. Howie ruled that the conditions were invalid on all three grounds. He described the conditions as an "intolerable interference" with a university council's duty to have disciplinary powers administered freely and fairly.

Natal University's principal, Professor P. Booysen, said "I hope the matter will now be allowed to rest there and that the universities of South Africa be allowed to get on with the matter of regulating their own lives and their own communities, in such a way that we can meet our educational objectives".

Michael Cherry

Privatization poses questions for UK electricity research

London

BRITAIN'S electricity supply industry announced last week that it is to double its expenditure on research into possible health risks associated with magnetic fields produced by the distribution and use of electricity. The extended research programme, which for the coming year will increase to £1 million, will make use for the first time in Britain of newly developed personal exposure meters to help build up an accurate picture of magnetic fields experienced by people in their daily lives. A decade of research into the question has failed to arrive at a firm conclusion.

The Central Electricity Generating Board (CEGB) is at present supporting two independent research projects, at the universities of Leeds and Manchester and at health authorities in Yorkshire and Lancashire, looking for possible links between living near power lines and adult and childhood cancers. Preliminary results from the childhood study have found no association; the results from the adult study are expected next year.

A project commissioned by the New York State utilities and published last year examined possible biological effects of magnetic and electric fields. One study, in Denver, suggested a small positive correlation between childhood cancer and proximity to power lines. The new British study will involve volunteers from the electricity supply industry wearing personal exposure meters at home and at work; assessing magnetic fields in homes by using new vehicle-mounted sensing equipment; and a study by independent epidemiologists of all newly diagnosed childhood cancer cases in England and Wales that will enable the child's past exposure to be better estimated.

The announcement of the new research programme has, however, prompted renewed speculation about the fate of the industry's research interests after its privatization. Earlier this month, the government produced a white paper (policy document) outlining its plans for how the industry should be sold off. It is proposed that the 12 distribution organizations (area boards) should become private companies. The CEGB will be split into two parts, the larger of which will own the nuclear power stations. The white paper does not mention the fate of the industry's considerable research capability.

Last year the CEGB spent £162 million on research and development. Most of this (68 per cent) went on nuclear research, while 16 per cent was spent on research into the environment and new energy sources. The CEGB's Technology



Monitoring magnetic fields generated by power lines. But would future shareholders approve?

and Planning Research Division employs 800 people, mainly at its three laboratories. Precisely who will have responsibility for research activities after privatization is a question that is worrying CEGB researchers, as is the question of the new company's level of commitment to non-profit-linked research. The CEGB has already pulled out of one joint project, into the decommissioning of nuclear reactors.

The board is examining ways of organizing its research after privatization, but few individual scientists seem to know what is to become of them. As an employee at CEGB's Central Electricity Research Laboratory in Surrey speculated last week, "We're on a prime site next to the M25, so we'll probably be turned into a supermarket".

Simon Hadlington

UK's biotechnology lacking specialists

London

BRITAIN will not be able to keep pace with its industrial competitors in the expanding field of biotechnology unless serious manpower and training problems are addressed. Such is the chief implication of a survey by the Association for the Advancement of British Biotechnology of its 80 member companies, whose results were released last week.

Shortages of manpower are reported both at the first degree and PhD levels in fermentation, protein chemistry and plant molecular biology, and less so in enzymology and chemical engineering. The shortage of plant molecular biologists, due to the increasing emphasis on plant biotechnology, is of particular concern to the industry. Furthermore, the recent trend to concentrate efforts on plant gene technology has slowed down advances in fundamental plant biochemistry and physiology. Efforts are being made to redress the balance, notably by a joint government-industry programme launched last year to encourage research into the control of plant metabolism.

The study predicts a rapid expansion of application in protein engineering, with a consequent shortage in enzyme technology, protein chemistry and biochemical engineering. Recognition of the importance of training in management, business administration and marketing is increasing, the study says, but only gradually compared with Britain's main trade competitors.

Simon Hadlington

High priority no guarantee for Germany's graduate colleges

Munich

WEST Germany has announced a scheme to improve graduate education at universities. The plan, introduced on 22 February by the influential Wissenschaftsrat (science council), would establish 80 'graduate colleges' for the systematic education of doctoral candidates. The West German government and the *Länder* (states) are being asked to finance the new colleges with up to DM40 million annually.

Sixteen graduate colleges exist already in fields such as molecular biology, engineering sciences and history of science. The colleges organize university faculty members to give courses to the students and allow each student to become familiar with a number of subfields. Students who have completed their *Diplom* (the rough equivalent of a master's degree) in close to the minimum of eight or nine semesters are favoured in the competition for scholarships. In addition to lowering the

average age of doctoral recipients, the colleges help to soften the rigid system of one-on-one graduate advising that has become the norm.

Encouraged by the programme's success, the Wissenschaftsrat wants the government to take the lead in expanding it. Ten of the existing colleges were established by charitable foundations such as Stiftung-Volkswagenwerk and six more by the federal and *Länder* governments on an experimental basis.

The lion's share of new federal support would fund up to 1,200 doctoral and 160 postdoctoral scholarships annually. The *Länder* would have to pay material and running costs for the colleges.

The Education Ministry has given the programme a "very high priority", according to a spokesman, but its success cannot be guaranteed in the forthcoming negotiations with the hard-pressed Finance Ministry.

Steven Dickman

The Shroud of Turin

SIR—Both the scientists involved and outside observers such as myself have been astonished at the recent decision of the Archbishop of Turin, Anastasio Ballestrero, to withdraw from four of the seven participating laboratories permission to carbon-test the Shroud of Turin.

This action, supposedly made in the interests of conservation of the shroud linen, leaves in a shambles the carefully devised plans of the group of experts who met in the autumn of 1986 to draw up testing procedures for the cloth. As things now stand, only laboratories at the University of Arizona, the Technical University in Zurich and the University of Oxford will be given shroud samples. Shut out from the tests will be Dr Harry Gove of the University of Rochester and Dr Garman Harbottle of the Brookhaven National Laboratory, as well as the Saclay Laboratory in France and the Atomic Energy Research Authority in Harwell.

Of equal importance is the fact that the Vatican officials in charge of the test have still not come forward with procedures to secure the authenticity of the samples themselves—procedures, for example, to make it impossible for ancient mummy linen to be surreptitiously introduced into the chain of evidence. If the shroud linen is itself of ancient origin, but the tested samples are not provably from the shroud, then there will be no reason for anybody to take the test results seriously.

I call on all the concerned laboratories to withdraw from the tests until such time as the Vatican decides to go back to the seven-laboratory plan, with strict, open procedures to ensure the authenticity of samples.

DENIS DUTTON

*School of Fine Arts,
University of Canterbury,
Christchurch, New Zealand*

SIR—Those in a position to apply scientific tests to the Shroud of Turin might care to look for asbestos fibres in it, as well as in the other folds of the shroud reported (by Murray's *Guide to Northern Italy*, 1883) to exist in Rome, Besançon and Cadouin.

Marco Polo, after describing the extraction of asbestos in China, from mines perhaps near Hami in Sinkiang, says that the fibres, which look like wool, are spun into napkins which may be cleaned by being put into the fire. He continues: "And I will also tell you that in Rome there is a napkin that the Great Kaan sent to the Pope as a splendid present, when he sent the two Polo brothers [Niccolo, Marco's father and Matteo his uncle] to him as envoys, in order that the sacred shroud of Our Lord Jesus Christ might be wrapped up in it. And on that napkin are

golden letters, saying Tu es Petrus et super hanc petram edificabo ecclesiam meam" (*Travels of Marco Polo*, Routledge & Kegan Paul, 1931, p.74).

In fact, in 1269, when the Polo brothers arrived back with this present from Kublai Khan, it was in a period (1268–71) when there was no Pope. The Shroud of Turin, apparently, was not reported before the fourteenth century, but we may wonder whatever happened to the fireproof cover, especially as some have considered that the shroud carries scorch marks.

ALAN L. MACKAY

*Department of Crystallography,
Birkbeck College,
Malet Street, London WC1E 7HX, UK*

Independent Estonia

SIR—Contrary to the statement in your article "The Finn red line" (*Nature* 330, 331; 1987), Estonia was not part of Finland before the Second World War but was an independent republic in its own right.

Who's rewriting history now?!

T. OJASOO

*17 York Road,
Stafford ST17 4PE, UK*

Moral failure?

SIR—Why is the AIDS pandemic accompanied by paralysis of the faculty of logical reasoning as none other? The leading article "Whose own goal?" (*Nature* 331, 376; 1988) is a case in point. It concludes "AIDS is not a moral failure but a viral infection". Why should the two be mutually exclusive?

For the sake of argument let us consider that there is an epidemic of red noses. Let us first suppose that red noses are caused by the transmission of a virus from one individual to another by, say, rubbing noses. Second, let us suppose that this is more likely to happen when males rub noses (homonal contact) in Europe, but far more readily transmitted by heteronasal contact in Africa. Transmission by blood is unfortunately also becoming more frequent. What public health measures need to be instituted to contain the pandemic? First, nose rubbing needs to be avoided except with one partner not infected with the virus. This is the 'no-risk option'. Secondly, and a flawed alternative, noses must be protected during multiple contacts. This is a 'lower-risk option'. Blood transmission needs to be dealt with separately.

In this scenario, alteration in human conduct is necessary for the greater good of all. One of the *Oxford English Dictionary's* definitions of morality is "principles dealing with regulation of conduct". Prevention of infected nose rubbing is, there-

fore, to do with morality and public health.

Thus it is with AIDS—a viral pandemic produced by patterns of human behaviour which deviate from a fundamental code of moral behaviour. The further tragedy is that innocent parties are becoming increasingly and lethally vulnerable.

The Princes's Royal was absolutely correct in what she said. AIDS is indeed "mankind's own goal"—a viral infection turned into a pandemic by moral failure.

GORDON M. STIRRAT

*University of Bristol,
Department of Obstetrics
and Gynaecology,
Bristol Maternity Hospital,
Southwell Street, Bristol BS2 8EG, UK*

Arabs and Israelis

SIR—I was extremely surprised by Nechemia Meyers' article entitled "Arab-Israeli cooperation goes public" (*Nature* 330, 302; 1987).

I have never been involved in any "cooperative science project" with Israeli researchers. My only contacts with researchers from Israel have been seminars or scientific meetings organized by the French Cancer Research Association (ARC) or international organizations such as the UICC or WHO. The second day of a two-day ARC-sponsored French-Mediterranean colloquium, on 17 November 1987, was used for the presentation of the "Preliminary Results of the Tunisian-French Research Project on Inflammatory Breast Cancer". As the person responsible for this project, I was asked to be moderator of the one-day discussion on inflammatory breast cancer. Dr Feldman, on the other hand, was responsible for the organization of the other day, which treated a subject with absolutely no relation to mine. It is therefore difficult to see how, from such superficial contact, "the potential for collaboration between Arab and Israeli researchers" could be considered as "evident".

Let me conclude by saying that I am not only a government officer, but am also in total support of the position held by my government: I believe that the most important question at the present time is not scientific cooperation, but the search for a political solution to the Palestinian problem.

NEJIB MOURALI

*Institut Salah Azaiz,
Bab Saadoun,
Tunis,
Tunisia*

● NECHEMIA MEYERS ADDS: I had not intended to imply that Professor Mourali had been involved in a cooperative research project but only that there is potential for collaboration that will materialize only, as I quoted Professor Feldman as saying, if "politicians get their act together". □

Making the geoid respectable again

With the recognition that surface measurements of gravity cannot be extrapolated backwards into the Earth's interior, the concept of the geoid became unfashionable. But its fortunes may be on the turn.

EVERYBODY knows, of course, that the geoid is that surface of equal gravitational potential which coincides with mean sea-level. Most of us also have it clearly in mind that the geoid is an ellipsoid with rotational symmetry about its shorter polar axis, which allows for the flattening of the Earth by rotation. But that must plainly be a fairy story, possibly applicable to the oceanic surfaces of the Earth, but unlikely to make very much sense elsewhere. The snag, as can be learned from a glance at a recent article by R.G. Hipkin from the University of Edinburgh (*Geophysical Journal* **92**, 53; 1988), is that it is much easier to stick to the fairy-story than to tell just where the geoid is on dry land.

Hipkin's opening sentence may be a sufficient explanation:

Physical geodesy has become a dauntingly esoteric branch of applied mathematics which, in the achievement of greater rigour, has distanced itself from many practical geodesists. This is particularly unfortunate when greatly improved measurement techniques have become available, both for surface gravimetry and for geometrical positioning from space vehicles.

Hipkin might have added that geodesy may not be the only offender in this regard.

Whatever the validity of that *canard*, there is clearly plenty of room for intricate argument about even the elementary definitions on which geodesy is founded. Hipkin's approach is refreshingly that of a logical positivist of the machian school. In pursuit of the principle that the only quantities appearing in equations should be measurable quantities, he even rejects Newton's scheme (really, a *gedanken-experiment*) for defining the geoid everywhere by building a series of wells and canals within the body of the Earth; the construction would vitiate the observations, he says.

With that said, Hipkin's interest is to rescue the concept of the geoid from the limbo into which it may be expelled by excessive rigour. (With tongue firmly in cheek, he quotes two writers from twenty years ago who say that this ambition is a "concession to conventional conceptions"). But the difficulties are considerable, as Hipkin's article shows clearly enough.

Were it not for the continents, and anomalies of gravitational potential such as that in the south-west Atlantic or the non-circularity of the Equator, the surface

of mean sea level would indeed be an ellipsoid of rotation. One of the practical difficulties with which geodesists are confronted is that the measurements of sea level at tidal stations must be corrected for such sources of variation as ocean currents, departures of seawater density from the means, not to mention other seasonal effects, but in principle there is a wealth of places scattered around the world at which the ellipsoidal parts of the geoid might be defined precisely, especially now that satellite ranging data may provide locations more precisely (or at least more conveniently) than by timing stars.

One superficial complication is that the reference ellipsoid and the set of concentric ellipsoids corresponding to different values of the gravitational potential define a downward gravitational attraction (called "normal gravity" and calculated as $\gamma = \nabla U$, where ∇ is the laplacian operator and U is the potential field) that will differ both in magnitude and in direction from true gravity, g (notionally calculable from the true gravitational potential W as $0.179W$).

Hipkin does not differ from others in remarking that it is in principle possible to relate measurements of gravity (including the departure of its direction from that of the line to the centre of the Earth) at inland surface sites to measurements at coastal tidal stations by a combination of levelling and gravity measurements (but care is needed to distinguish between the normals to the ellipsoidal surfaces and the direction of a plumb line).

The essence of the difficulty is that of telling what happens to the force of gravity beneath the surface, within the continental crust for example. Formally, nothing can be done without a knowledge of the density distribution of material within the body of the Earth. While, as every school-boy knows (or, perhaps, should know), the external gravitational field of the Earth is uniquely determined by a knowledge of its value and gradient at each point on the surface, nothing can be said about the gravitational field within the Earth without a detailed knowledge of the density distribution within the solid Earth. That seems to be the starting point, apparently established by the Soviet mathematician Molodensky, for the recently rigorous trend in geodesy.

Plainly, there are also serious mathematical problems in extrapolating downwards from the surface of the Earth even

when there are ample measurements of gravity on which to base inferences. Hipkin gives the example of a ploughed field, the furrows of which would yield small but still measurable periodic variations of gravitational potential just above the surface which, if projected downwards in a simple-minded fashion, would imply variations of the height of the equipotential surface of as much as 1 km just a few metres below the surface.

Hipkin's goal is to find a sensible and consistent way of projecting surface measurements of gravity downwards. This he does by splitting the measured gravity anomalies at the surface into two parts — that due to variations from flatness of the terrain in the immediate neighbourhood and that represented by what is usually called the Bouguer anomaly, and which corresponds to deep-seated gravity anomalies such as the departures from hydrostatic balance, more precisely isostasy.

The proof of the recipe is, as always, whether it works. On the face of things, and for a careful set of data collected in northern Scotland, the answer is encouraging. First, it appears practicable to deal with the effects of topographic variations (of which there are many, called mountains, in northern Scotland) by making the simple assumption that they consist of material of constant density and then by estimating their contribution to local gravity by means of a contour-following technique. Second, the downward projection of the Bouguer anomaly itself can be arranged in a consistent fashion by careful control of the wavelength of the measured surface gravity field.

The result, for northern Scotland, is a more detailed account of where the geoid lies than any other part of the world is provided with. On the face of things, the height of the geoid there is defined with a precision of less than 10 cm. These variations become apparent after allowing for an east-west tilt in the surface of the geoid from one side of Scotland to the other. In principle, and in due course, this should provide grist for the mills of those wishing to speculate about the reasons why, for example, there should be a relative depression of the geoid amounting to a maximum of more than 1.5 m off the northern coast of Aberdeenshire. But the general interest of what Hipkin has done is to make the concept of the geoid respectable again.

John Maddox

Archaeology

Triple Czech burial

Paul G. Bahn

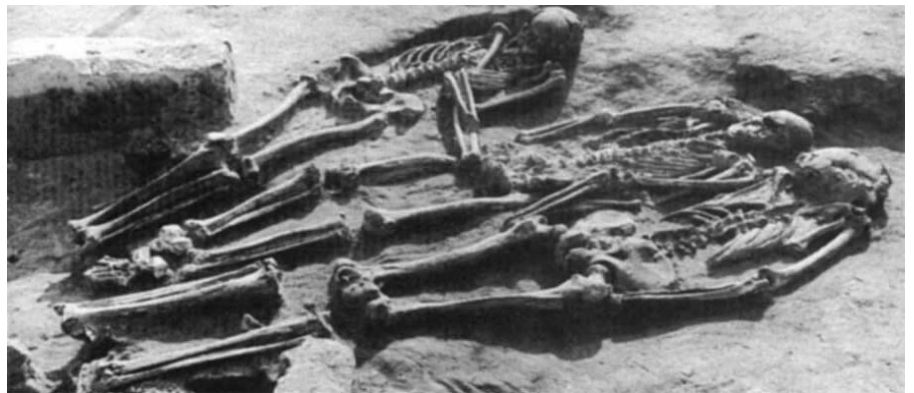
SINCE they were first discovered and authenticated in the late nineteenth century, Upper Palaeolithic burials have become quite numerous^{1,2}, although still very rare in the context of the period's length and the number of people who must have died. It seems likely that many of those accorded careful or elaborate burial in the last Ice Age were special. This supposition is strengthened by the recent report³ of a 17-year-old male dwarf buried during the late Upper Palaeolithic in an important decorated cave in Italy. A further Palaeolithic grave has now come to light^{4,5} during excavations by Bohuslav Klíma at the famous open-air site of Dolní Věstonice in the Pavlovské Kopce Hills of southern Moravia, where decades of work by researchers from the Czech Academy's Archaeological Institute have provided many insights into Early Upper Palaeolithic life — not just camps and burials but also the remarkable terracotta figurines and their 'kiln'.

Whatever the events and medical history behind this unique Palaeolithic grave, it adds to our picture of the complex nature of beliefs and funerary rituals in the last Ice Age. It is worth noting that, as in this case, isolated human bones have often been found around elaborate graves of the period, not only at

Dolní Věstonice and nearby Pavlov, but also at Předměstí and even further afield at Sungir, near Moscow.

A rescue excavation in 1986, in a level at Dolní Věstonice whose loess deposits were being exploited by heavy machinery, led to the discovery of another camp of the Pavlovian period. A whole line of hearths

pelvis. The three skulls were in a soil impregnated with red pigment (the bigger man's skull also had a coating of white powder), and the red crust around their frontals contained the remains of 'diadems' of wolf and fox teeth and small ivory beads: such items are fairly common in burials of this period, particularly in central and eastern Europe. There was also a considerable concentration of ochre under the female's pelvis and between her thighs. The pit contained a few small fragments of animal bone (reindeer and wolf), some flint flakes, an engraved pebble and a cluster of unperforated shells.



A mystery surrounds the deaths of three Ice Age skeletons buried together over 26,000 years ago.

extended towards the hills, and contained large fragments of carbonized wood; there were also many tools of stone and bone. A radiocarbon date of $27,660 \pm 80$ years before present (BP) was obtained (GrN-13692) from this level.

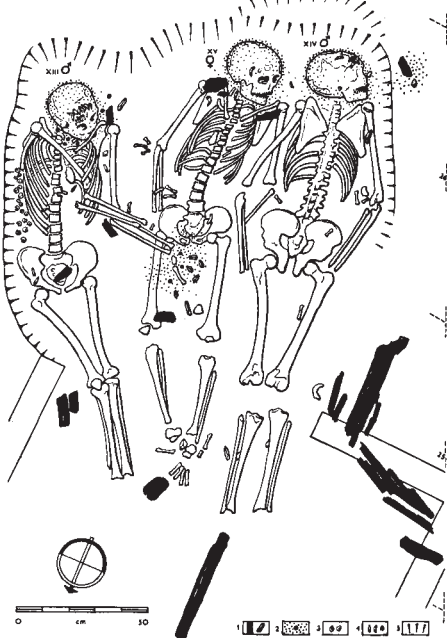
The upper part of a human skull (Dolní Věstonice XI) belonging to a man about 40 years old was found here near a hearth, together with bones of mammoth, reindeer and wolf. A fragment of frontal bone (DV XII), probably from the same individual, lay 15 m away, and displayed a large wound above the right eye which had healed but which probably had lasting effects on the individual's health. It seems to have been caused by a hard, blunt instrument wielded with some strength, and may thus represent one of the (very rare) clues about the occurrence of violence during the Upper Palaeolithic⁶.

On 13 August 1986, however, a remarkable triple burial was found only 15 m from the skull and 30 m from the frontal bone (see figures)^{4,5}. The skeletons were in a shallow pit, in an extended position, side by side, with their heads towards the south. All three were young, between 16 and 23 years of age. The youngest, lying supine in the centre (DV XV), was very gracile; this fact, together with pelvic shape, points to a female. On her left was a big robust skeleton (DV XIV, a male) lying on its stomach, with its head turned away to the west; its left arm covered the female's hand, as if holding her. On her right was a smaller robust specimen (DV XIII), lying on its left side and facing the female; both of its arms touched her

A quantity of fragments of carbonized wood, up to 50 cm in length and 10 cm wide, lay directly on the skeletons as well as near them. These may be the remains of some sort of tomb-cover, a replacement for the covering of mammoth shoulder-blades usually found on Upper Palaeolithic graves in this area; it was perhaps burnt as part of the funerary ritual, as the greatest mass of carbonized wood (a 'funeral pyre'?) lay immediately west of the grave. The corpses themselves were not burned, but the presence of brick-red burnt loess immediately above the wood suggests that the fire was extinguished by having earth heaped over it. Wood from the grave has produced a radiocarbon date of $26,640 \pm 110$ years BP (GrN 14831).

The three skeletons had clearly been buried at the same time (the female was put in first) — are they there because of an accident, an epidemic, or for some other reason? One clue is that the female's bones show several pathological deformations. Her skull is strikingly asymmetrical and has a perforation from the frontal cavity to the left supraorbital torus; she had scoliosis of the vertebral column, deformation of the right iliac, and an insufficient development of the entire right side and limbs. These conditions suggest that she may have contracted rickets or encephalitis at a very early age. In her mouth was a fragment of burnt reindeer pelvis which could have been used as a clamp to bite on during times of great pain as it has traces of scratches and pressure.

One explanation, put forward by



Layout of the grave at Dolní Věstonice. (1) Charred wood fragments; (2) concentration of red ochre; (3) unperforated shells; (4) displaced human teeth (through the action of burrowing animals), and perforated wolf and fox teeth; (5) edge of grave. (Courtesy of B. Klíma.)

Bohuslav Klíma, is that the burial is a reproduction of a real event — a failed birth. The man on the female's left holds her arm to give comfort, but cannot bear to look, while the man on her right tries to help the delivery, but in vain. Klíma suggests that the ochre around the female's pelvis may mark the original location of a fourth skull or the complete skeleton of a newborn child whose bones have disintegrated.

The idea is appealing, but the supposedly complete disappearance of the newborn baby's remains is made slightly doubtful by the good preservation and virtually complete state of the three skeletons. In addition, there are other enigmas: for example, the male with his hand on the female's pelvis was skewered to his sacrum by a large piece of wood, while the other male's skull was smashed (and further deformed by the weight of earth

on it). Research is under way to test the hypothesis that these two apparently healthy and powerful men met violent deaths, perhaps sacrificed to accompany the dead female. Moreover, this handicapped female, like the dwarf², the crippled old Neanderthaler from Shanidar, Iraq⁷ and others, adds further proof to the already well-established notion that communities of the Middle and Upper Palaeolithic took good care of — and perhaps even accorded special importance to — physically disadvantaged individuals. □

1. Harrold, F.B. *World Archaeol.* **12**, 195 (1980).
2. Binford, S.R. *Southwestern J. Anthropol.* **24**, 139 (1968).
3. Frayer, D.W. *et al. Nature* **330**, 60 (1987).
4. Klíma, B. *Quartär* **37/38**, 53 (1987).
5. Klíma, B. *L'Anthropologie* **91**, 329 (1987).
6. Roper, M.K. *Curr. Anthropol.* **10**, 427 (1969).
7. Trinkaus, E. *The Shanidar Neanderthals* (Academic, New York, 1983).

Paul G. Bahn is a freelance archaeological writer at 428 Anlaby Road, Hull HU3 6QP, UK.

Protein structure

An extra dimension to NMR

Dagmar Ringe

TWO-DIMENSIONAL nuclear magnetic resonance (2D-NMR) techniques, developed by Ernst and Wüthrich^{1,2}, have greatly advanced the determination of protein structures in solution. The geometric information that is needed to solve the three-dimensional structure of a protein by NMR spectroscopy is obtained from interproton distances assigned from correlation and nuclear Overhauser effect (NOE) experiments. The quality of the structure obtained depends critically on how well the extensive spectral overlap that is observed can be resolved so as to enable the resonances to be assigned unambiguously to specific protons. On page 374 of this issue, Oschkinat *et al.*³ describe a technique of further improving the resolution of the cross-peaks by extending 2D-NMR into a third dimension.

NMR is proving to be an important complement to X-ray crystallography in the determination of protein structure. Although there is as yet no concrete evidence to indicate that the structures of proteins in solution, determined by NMR, are substantially different from those of proteins in crystal form, determined by X-ray diffraction, provided both techniques have been applied correctly, there is, in principle, the possibility that the restraints imposed on a protein structure by the crystal environment could affect the tertiary structure of segments of the polypeptide chain. In addition, although water of crystallization is observed as a bound surface layer in protein crystals, structural information obtained crystallographically cannot give any extensive information about the effect of bulk solvent on the

conformations and mobilities of many surface residues of a protein. Furthermore, although some information on the spatial distribution of the dynamics of a protein can be obtained by X-ray crystallography⁴, such information is practically intrinsic to NMR spectroscopy and is therefore more readily accessible. And finally, the range of sizes of polypeptides and proteins for which NMR data are most obtainable lies in precisely that region for which X-ray structures are often hard, or impossible, to obtain because small proteins are difficult to crystallize, present few reactive surface residues for derivatization with heavy metals, and are too large for phasing by direct methods. Application of NMR spectroscopy to larger proteins has developed slowly; the effective upper limit is still 10 kilodaltons in the absence of X-ray crystallographic information to help make assignments.

A protein structure is obtained from NMR data by applying interproton distance information to a distance geometry algorithm⁵ or to a molecular dynamics simulation starting from the unfolded polypeptide chain⁶. Because the NOE data required for the solution of a three-dimensional structure by NMR can only give information on short-distance interactions (less than about 5 Å), it becomes critically important to be able to resolve as many of the peaks arising from such interactions as possible. In general, the resolution of the cross-peaks arising from interproton interactions in a β -sheet is relatively straightforward. But cross-peaks that arise from interactions between elements of secondary structure,

particularly those arising from interactions of less well-defined secondary structure such as loops or from α -helices, are very often not well resolved, partly because they are weak and partly because they overlap cross-peaks from other interactions. Although the technique that Oschkinat *et al.* describe cannot improve the former, it can significantly improve the latter.

The authors have chosen α 1-purothionin, a protein of 46 residues, to illustrate the utility of their technique. In particular, a NOE spectroscopy experiment (detection of inter-residue and intra-residue NOE connectivities between NH and C¹H protons) was combined with a total-correlated spectroscopy experiment (detection of J-coupling of intra-residue connectivities between C¹H(i) and NH(i) protons) in a restricted region of the NMR spectrum. Recording a complete 3D-NMR spectrum is not possible because it would take too long. But the technique can be applied selectively in regions with extensive overlap of cross-peaks. Thus, the combined experiments described in the paper allowed the assignment of inter-residue cross-peaks that could not be assigned uniquely by 2D-NMR, owing to overlap with other cross-peaks.

Another means to extend resolution has recently been reported, with thioredoxin (108 amino-acid residues) as the example⁷. The thioredoxin was randomly labelled with deuterium to the extent of approximately 75 per cent. The resulting protein gives a significantly better resolved 2D-NMR spectrum than the protein with natural abundance hydrogens. It is conceivable that a combination of the isotope labelling method and extension to 3D-NMR will make it possible to tackle proteins substantially above the 10 kilodalton range, including many that carry out interesting reactions or that undergo conformational changes which cannot be studied by crystallographic means. The only drawback of the three-dimensional method would seem to be that readers of journals and those attending conferences, who have been confounded by the incomprehensible 2D-NMR plots so dear to the hearts of NMR spectroscopists, will now be even further confounded by stereo diagrams of three-dimensional arrays of spots. □

1. Ave, W.P., Barthold, E. & Ernst, G.G. *J. chem. Phys.* **64**, 2229–2246 (1976).
2. Kline, A.D., Braun, W. & Wüthrich, K. *J. molec. Biol.* **189**, 377 (1986).
3. Oschkinat, H. *et al. Nature* **332**, 374–376 (1988).
4. Ringe, D. & Petsko, G.A. *Prog. Biophys. molec. Biol.* **45**, 197–235 (1985).
5. Crippen, G.M. *J. comp. Phys.* **24**, 96–107 (1977).
6. Brünger, A.T. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 3801–3805 (1986).
7. LeMaster, D.M. & Richards, F.M. *Biochemistry* **27**, 142–150 (1988).

Dagmar Ringe is in the Chemistry Department at the Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

Conservation biology

Red books or green lists?

Jared M. Diamond

A RECENT article¹ by Christoph Imboden, director of the International Council for Bird Preservation, proposes a strategy to correct a major tactical error on the part of conservation biologists. For the past 25 years the status of the world's biota has been summarized in so-called Red Data Books. These volumes list species definitely known to be threatened, thereby implying that all remaining species are secure. Imboden now suggests "green lists" to contain those species definitely known to be secure, thereby shifting the burden of proof to those who maintain that all is well. Several recent studies illustrate how green lists will give a much more realistic indication of the extinction crisis than the red books.

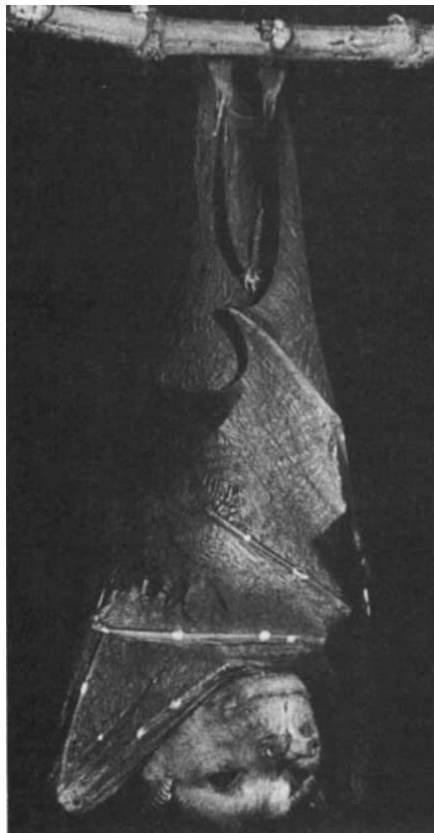
The mentality behind the red books is appropriate to Europe and North America, where armies of amateur naturalists search for rare birds every weekend, and where field guides advise how to find and recognize rare species. Thus, up-to-date information is always available on the conservation status of every species. It could not be that a European or North American bird species would suddenly become extinct without people having known for years that it was endangered. The last wild California condors were even identified individually. For such well-studied biotas the Red Data Books were maximal lists of endangered species.

This approach, applied around the globe, led to the conclusion that 87 of the approximately 8,000 bird species extant in the year 1600 are now extinct, and that about 283 more are endangered². Advocates of unlimited economic growth dispute the evidence in some of those 283 cases, and note with satisfaction that even conservationists consider $(100)/(8,000 - 283 - 87)/(8,000 - 87) = 96$ per cent of the world's bird species are secure, and trumpet that all is well.

But most species live in the tropics, where there are few field guides to tell the few bird-watchers where to look for a rare species or how to identify it even if they should find it. One cannot assume that lack of bad news about a species guarantees good news, as it does for European and North American birds. Nothing is known about the current status of many tropical species discovered long ago, simply because no biologist has visited the collecting locality since the original description. Other tropical species are known only from a few nineteenth-century specimens taken at an unknown locality^{3,4}. Given the small

geographical ranges often characterizing tropical species, many must have vanished in formerly forested areas now converted to croplands and pastures. Hence Red Data Books for the tropics are minimal rather than maximal lists of endangered species, and only case-by-case assessments can show which of the remaining species actually belong on a green list.

Such assessments became available within the past year for Sangihe and



The tube-nosed fruit bat Nyctimene rabori, confined to one or two forest fragments on the island of Negros, undescribed until 1983, and already endangered by forest destruction.

Talaud island birds, Solomon Island birds and Philippine fruit bats. In the nineteenth century, five species of endemic birds were discovered by expeditions to the Sangihe and Talaud islands of Indonesia, whereupon the islands relapsed into ornithological obscurity. When ornithologists finally revisited the islands in the 1980s, it turned out that much of the forest had been felled for plantations, and in the remaining forest patches only one pair of the Sangihe hanging parrot, one male of the Sangihe elegant sunbird, and no Caerulean paradise flycatchers were to be found⁵. Recent surveys of the Solomon Islands failed to re-encounter

12 of the 164 previously described bird species on the archipelago. Eight of the missing twelve species were ground dwellers, of which one is known and the other seven are likely to have been exterminated by introduced cats⁶. Philippine fruit bats include the world's heaviest bat (crested fruit bat, 2.5 lb) as well as the bat with the widest wingspan (large flying fox, 6 feet). Roost sizes of both these giants crashed from more than 100,000 in the 1920s to a few hundred individuals in the 1980s; the shining flying fox has never been seen again since its discovery in 1888; the status of three other large species is virtually unknown; Chapman's bare-backed fruit bat, discovered in 1952, was extinct by 1981; the tube-nosed fruit bat (see figure), which remained undescribed till 1983, is now confined to a shrinking fragment of forest; and only two species are known to be still common⁷.

A striking fact about almost all these Indonesian, Solomon Island and Philippine species is that they were not listed in the red books: they became extinct, or reached the verge of extinction, before biologists realized their peril. In other recent cases the interval at which extinction followed realization of danger or even of a species' mere existence was barely a few years. Bulmer's fruit bat, described as a fossil species thought to have become extinct 10,000 years ago, was discovered alive in the mountains of New Guinea in 1975, only to be "re-exterminated" by hunters within the next two years⁸. In 1976 it was realized that there had been no sightings of the formerly abundant sheath-tailed bat of the Marianas since 1972. An ironical decision by the US Fish and Wildlife Service then refused to list the species as endangered, on the grounds that there was no evidence for its continued existence! The rediscovery of a colony of less than 10 in 1984 now starts a new race: will this population too be exterminated before its endangered status is officially acknowledged⁹?

The real test of whether red books or green lists give a truer picture of the extinction crisis will come from the world's most species-rich habitats, the continental tropics. There are innumerable species of tropical Asia, Africa and South America for which we have no recent information and, for which there have been no recent specific searches,

1. Imboden, C. *World Birdwatch* 9 (2), 2 (1987).
2. King, W.B. *Endangered Birds of the World* (Smithsonian Instn, Washington DC, 1981).
3. Terborgh, J. & Weske, J.S. *Auk* 89, 497-505 (1972).
4. Olson, S.L. *Le Gerfaut* 76, 263-269 (1986).
5. Whitten, A.J. *et al. Conserv. Biol.* 1, 42-48 (1987).
6. Diamond, J.M. *Conserv. Biol.* 1, 77-79 (1987).
7. Heaney, L.R. & Heideman, P.D. *Bats* 5, 3-5 (1987).
8. Hyndman, D. & Menzies, J.I. *J. Mammal.* 61, 159 (1980).
9. Lemke, T.O. *Bats* 5, 5-7 (1987).
10. Mohsin, A.K.M. & Ambak, M.A. *Freshwater Fishes of Peninsular Malaysia* (University Pertanian Malaysia Press, Kuala Lumpur, 1983).
11. Olson, S.L. in *Conservation Biology for the Next Century* (eds Pearl, M. & Western, D.) (Oxford University Press, in the press).

which are not mentioned in the red books.

As one indication of what such surveys may reveal, four years of searching (1979–83) for primary freshwater fishes on the Malay peninsula, where 266 species had at one time been found, uncovered no evidence for the continued existence of an astonishing 54 per cent of those species¹⁰. Imboden's estimate¹ is that fewer than one-third of the world's bird species will qualify for a green list, defined by the criterion of "known not to be declining in numbers now, and unlikely to decline in

the next decade". Recent archaeological discoveries indicate that up to one-quarter of the bird species that existed a few millennia ago have been exterminated in modern times, but before they could be described by ornithologists, when humans colonized oceanic islands¹¹. Thus, the much-discussed extinction crisis is not a debatable future possibility: it is already well advanced. □

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024, USA.

Cuprate superconductors

Structure and superstructure

Colin Greaves and Ted Forgan

PROGRESS in understanding the new rare-earth-free superconductors that we discussed in a recent News and Views article¹ has been rapid. We suggested then that the structure of Bi–Ca–Sr–Cu–O (BCSCO), which has superconducting transitions at 85 and 110 K, would soon be characterized. Several groups have now established that in the primary (85-K) phase, the metal ions are present in the approximate ratio 2:1:2:2. Veblen *et al.* report on page 334 of this issue² an electron-microscopy study that confirms the layered nature of this material, but clearly shows that the structure is anything but straightforward, as the basic building blocks seem able to fit together in slightly different ways, causing some disordering.

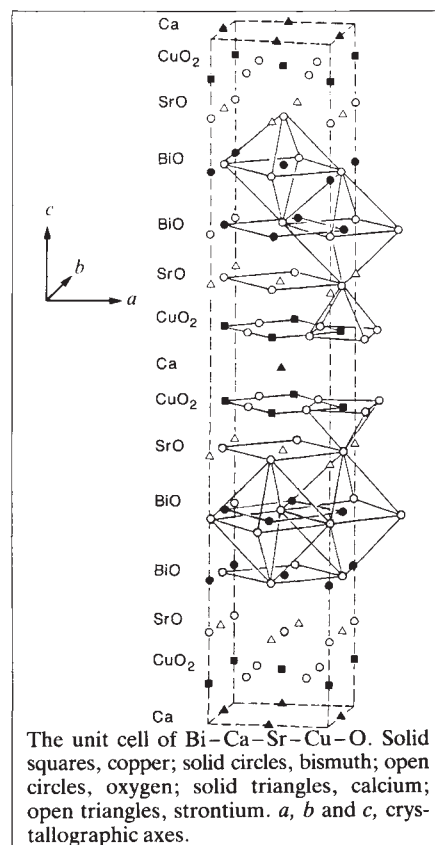
X-ray and neutron-diffraction studies of crystals with nominal composition $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_{8+x}$ (ref. 3; S.A. Sunshine *et al.*, submitted to *Phys. Rev. Lett.*; J.M. Tarascon *et al.*, submitted to *Phys. Rev.*) reveal the same pseudo-tetragonal structure ($a=b=5.4$ Å; $c=30.8$ Å; see figure) for the basic building block that Veblen *et al.* find. Although structural relationships between this material and the Aurivillius phases⁴ are often emphasized, it is clear that two important differences exist.

First, Bi_2O_2 layers, which are a characteristic feature of the Aurivillius phases and actually consist of a Bi–O₂–Bi sandwich, are absent. In the superconducting phase, the central oxygen atoms of the sandwich have moved into the bismuth layers to form two BiO layers with a rock-salt (NaCl) type of structure. These layers are quite weakly bonded together to form the cleavage planes of this material⁵. Second, the perovskite-like structure, which interleaves the Bi_2O_2 layers in the Aurivillius phases, does not begin immediately adjacent to these layers in BCSCO. Instead, the rock-salt arrangement is extended on both sides by additional layers of SrO type.

Between the rock-salt structural blocks there are two CuO_2 layers separated by

calcium (Ca^{2+}) ions. Each copper atom has four coplanar oxygen neighbours. Another more distant oxygen atom in the adjacent SrO plane completes a square-pyramidal arrangement. The double copper layers are very similar to those found in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$, in which Y^{3+} rather than Ca^{2+} is the bridging cation. These layers, therefore, increasingly seem to be important in the new superconductors, because the bismuth-containing material contains no copper chains analogous to those in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. We expect that the electrons responsible for the superconductivity belong to copper planes, and the copper ions have an average valence between two and three. It is also likely that the BiO planes are essentially insulating, so it would be no surprise if this material is even more two-dimensional than $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ (see ref. 6). It is interesting to note the pronounced structural similarity between this new phase and the layer materials described in 1980 by Nguyen *et al.*⁷, for example $\text{La}_2\text{SrCu}_2\text{O}_6$, in which rock-salt blocks (La_2O_2) separate CuO_2 –Sr– CuO_2 double layers.

The high-resolution micrographs reported by Veblen *et al.* in this issue² reveal departures from the idealized structure shown in the figure. First, there is disorder along the c -direction (perpendicular to the planes), which is not unexpected in layer materials of this type, and seems to correspond to the exclusion or introduction of an extra unit about 4-Å thick. Along the b -direction, the structural features have a much greater length scale than the 5.4-Å repeat distance of the basic building block and are related to a commensurate or incommensurate superstructure with a lattice spacing five times larger^{3,5} (27 Å). This could be related to the unusual bismuth coordination in BCSCO. Although the oxygen atoms around this site form a slightly distorted octahedron, the lone pair of electrons on bismuth (Bi^{3+}) ions generally results in much less symmetrical environments. It is possible that distortions that



lower the site symmetry could lead to buckling of the BiO layers to form a superstructure associated with periodic displacements. The satellite peaks seen on either side of the principal peaks along the b -direction in the electron-diffraction pattern² seem compatible with such distortions, which could also reduce the local symmetry to monoclinic, as Veblen *et al.* suggest.

Finally, some words of caution concerning the deduction of microstructural features from electron micrographs. The samples examined by Veblen *et al.*² contain traces of the 110-K superconducting phase, which is formed only by heating the 85-K phase close to its melting point. Specimens prepared at lower temperatures may not show the same type or extent of structural disorder². Also, the specimens were prepared by finely grinding a sintered sample, and some disorder could have been introduced by this process. If this is so, it is itself an important observation, as any practical applications of these materials may depend on their mechanical strength and structural integrity. □

1. Forgan, E.M. & Greaves, C. *Nature* **332**, 14–15 (1988).
2. Veblen, D.R. *et al.* *Nature* **332**, 334–337 (1988).
3. Bordet, P. *et al.* *Proc. Int. Conf. High-Temp. Superconductors Mater. Mech. Superconductivity* (in the press).
4. Aurivillius, B. *Arkiv. Kemi.* **1**, 463 (1949).
5. Hewat, E.A. *et al.* *Proc. Int. Conf. High-Temp. Superconductors Mater. Mech. Superconductivity* (in the press).
6. Dinger, T.R. & Tozer, S.W. *Nature* **332**, 204 (1988).
7. Nguyen, N. *et al.* *Mater. Res. Bull.* **15**, 891–897 (1980).

Colin Greaves is in the Department of Chemistry and Ted Forgan is in the Department of Physics at the University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

Immunology

Global or directed exocytosis?

Michail V. Sitkovsky and William E. Paul

T LYMPHOCYTES produce an array of exceedingly potent proteins called lymphokines that regulate almost every aspect of immune function¹. The purification, molecular cloning and large-scale production of many of these lymphokines could provide new pharmacological approaches to controlling the growth and action of lymphocytes and other blood cells. But the potential value of T-cell-derived lymphokines as drugs should not obscure the question of how they function physiologically. In particular, it is important to understand whether lymphokines are made by T cells for action at a distance (like endocrine hormones) or for action locally. The paper by Poo, Conrad and Janeway on page 378 of this issue² throws some light on this question.

T cells are generally activated and stimulated to produce lymphokines as a result of receptor-mediated interactions with cells that bear antigen-derived peptides bound to major histocompatibility complex (MHC) molecules^{3,4}. The very cell that presents antigen to the T cell and activates it to synthesize lymphokines is a potential target for the action of the T-cell product. This suggests that local accumulation of lymphokine, perhaps as a result of secretion into a limited space created by the apposition of the T cell and the antigen-presenting cell, is responsible for the specificity and precision of action of these antigen-non-specific products. Such a mechanism would ensure activation and/or selective expansion of the antigen-presenting cell, particularly when it is an antigen-specific B lymphocyte. High local concentrations of lymphokine could be achieved most efficiently if directed exocytosis causes its preferential secretion from the region of the T-cell membrane that contacts the antigen-presenting cell, where antigen-binding receptors are clustered as a result of interaction with antigen-MHC complexes on the antigen-presenting cell. It is this type of directional

exocytosis, in this case of the lymphokine interleukin-4 by a cloned helper T-lymphocyte line, that Poo *et al.* describe.

Directional exocytosis is a well-known phenomenon in secretory cells with polarized architecture, such as epithelial or neurosecretory cells. It has been suggested that it involves the specific association of secretory vesicles with cytoskeletal elements and the subsequent fusion of the vesicles with particular secretory domains of the plasma membrane^{5,6}. Non-polarized cells, such as mast cells and cytotoxic T lymphocytes, are thought to release granules towards the region of the cell membrane at which they were stimulated. Mast cells, for example, release granules containing enzymes and vasoactive amines close to the membrane sites at which their Fcε receptors (receptors for IgE) have been cross-linked⁷. These cells can be regarded as exhibiting transiently polarized secretion as a result of a newly applied activation stimulus. So far, only storage granules and their preformed contents have been shown to be directionally secreted in such cells.

Poo *et al.* now report² that interleukin-4 is directionally secreted towards the site of T-cell receptor occupancy. This result implies that focused secretion of newly synthesized proteins can also be achieved in transiently polarized T cells. Poo *et al.* extend earlier studies^{8,9} showing that cytotoxic T lymphocytes and natural killer cells reorient their microtubule-organizing centres and Golgi apparatus towards the area of contact between the cytotoxic cell and its target, suggesting that directed exocytosis is important in the cytolytic process. More recently, it has been shown that T cells of the helper phenotype (indeed, cells of the same line, D10, used in the experiments of Poo *et al.*) undergo similar reorientation of the microtubule-organizing centres and Golgi apparatus towards the site of interaction with antigen-presenting B lymphoma cells¹⁰. Although the role of these organelles in secretion has been questioned¹¹, their reorganization suggests that directional exocytosis does occur in transiently polarized T lymphocytes.

Poo *et al.* devised an assay which enabled them to examine directly whether D10 cells directionally secrete lymphokines towards the site of T-cell receptor occupancy. D10 cells are specific for conalbumin in association with class II MHC molecules of the I-A^k type and can be stimulated by conalbumin-bearing I-A^k-positive antigen-presenting cells or by soluble anti-T-cell receptor antibodies. Poo *et al.*

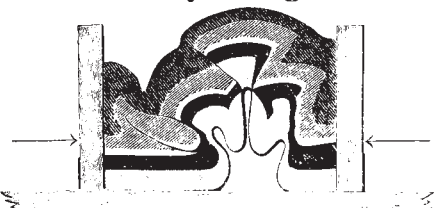
divided a culture chamber into upper and lower portions by hydrostatically plugging a nucleopore membrane with D10 cells, and introduced monoclonal antibodies specific for the D10 T-cell receptor into one chamber to obtain receptor occupancy mainly on either the upper or the lower surfaces of the T-cell membrane. They show that when anti-T-cell receptor antibody is added at low concentrations, interleukin-4 is preferentially secreted into the chamber in which the anti-receptor antibody is present. These results strongly suggest that T cells exhibit directed exocytosis towards the site of receptor occupancy by anti-receptor antibody and, by inference, to sites of receptor occupancy by complexes of antigen and MHC molecules. Therefore, the T-cell receptor (or associated molecules) can control not only lymphokine transcription but also the secretory process and, in particular, the site of exocytosis on the T-lymphocyte plasma membrane.

Thus, T cells and antigen-presenting cells can form transient, mobile synapses for the exchange of critical information. Such information exchange is probably bidirectional, as the antigen-presenting cell is stimulated by T-cell-derived lymphokines, and as T-cell activation depends on products of the antigen-presenting cell, such as interleukin-1 (ref. 12).

Focused exocytosis of interleukin-4 and related lymphokines into a 'synaptic' space between T cells and antigen-presenting cells is an attractive way to explain the specificity and selectivity of lymphokine function. The physiological significance of directional secretion of lymphokines is not yet understood. Is the interaction that triggers the T lymphocyte sufficiently durable so that the antigen-presenting cell that stimulates the T cell is also the beneficiary of the exocytic response? Do T cells exhibit multiple rounds of directed exocytosis towards the most recently occupied set of T-cell receptors as they sequentially encounter different antigen-presenting cells, or do they switch to a global secretion of lymphokine?

Finally, directed exocytosis does not account for all aspects of the function of T-cell-derived lymphokines. B cells that do not express appropriate complexes of antigen and MHC molecules on their surface can be activated when T cells are

100 years ago



The primary object of these experiments was to explain on what mechanical principles the remarkable structures recently discovered in the Highlands might have been produced. From *Nature* 37, 488; 22 March 1888.

1. Paul, W.E. *J. Immunol.* **139**, 1-6 (1987).
2. Poo, W.J., Conrad, L. & Janeway, C.A. Jr *Nature* **332**, 378-380 (1988).
3. Schwartz, R.H. *A. Rev. Immunol.* **3**, 237-261 (1985).
4. Babbitt, B. *et al.* *Nature* **317**, 359-361 (1985).
5. Kelly, R.B. *Science* **230**, 25-32 (1985).
6. Burgess, T.L. & Kelly, R.B. *A. Rev. Cell Biol.* **3**, 243 (1987).
7. Lawson, D. *et al.* *J. Cell Biol.* **79**, 394-400 (1978).
8. Geiger, B. *et al.* *J. Cell Biol.* **95**, 137-143 (1982).
9. Kupfer, A., Singer, S.J. & Dennert, G. *J. exp. Med.* **163**, 489-498 (1986).
10. Kupfer, A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 6080-6083 (1986).
11. Virtanen, I. *Eur. J. Cell Biol.* **42**, 281-287 (1986).
12. Durum, S.K., Schmidt, J.A. & Oppenheim, J. J. *A. Rev. Immunol.* **3**, 263-287 (1985).
13. De Franco, A., Ashwell, J.D., Schwartz, R.H. & Paul, W.E. *J. exp. Med.* **159**, 861-880 (1984).

specifically stimulated by antigen-presenting cells¹³. Such 'bystander activation' generally requires high concentrations of antigen. Bystander activation also appears to occur frequently *in vivo*. Indeed, this is consistent with the lack of complete polarity of interleukin-4 secretion observed by Poo *et al.* when they stimulated D10 cells with low concentrations of anti-T-cell receptor antibody, and the complete loss of polarity observed when they used high

concentrations of antibodies. Perhaps most importantly, their work emphasizes that the immune system is not a free-floating collection of independent cells, but exists as an organized entity, whose structure is critical for its regulated function. □

Michail V. Sitkovsky and William E. Paul are in the Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

Low-temperature physics

Superfluidity of ³He films

P. V. E. McClintock

THE superfluid flow properties of ³He films can now be studied in considerable detail, thanks to a new and highly sensitive technique developed at Berkeley by J. C. Davis, A. Amar, J. P. Pekola and R. E. Packard, described in *Physical Review Letters* (60, 302–304; 1988).

The original discovery of superfluidity, 50 years ago, arose in part through the astonishing observation by J. G. Daunt and K. Mendelssohn (*Nature* 141, 911–912; 1938) that any open-topped vessel of liquid helium, at a sufficiently low temperature, gradually empties. These authors established that a thin film of liquid creeps up the inside wall of the vessel, passes over its rim, slithers down the outside wall and then drips off the bottom. The process occurs because liquid ⁴He, below its transition temperature of 2.17 K, becomes superfluid; that is, it loses its viscosity and hence becomes able to flow rapidly through very tiny channels. One such channel is the 10–20-nm-thick adsorbed film of liquid that forms on the walls of the container under the influence of the van der Waals attraction.

Since 1972 it has been known that the rare isotope of helium, ³He, also undergoes a superfluid transition, and it is natural to ask whether it too can form a creeping superfluid film. The experimental difficulties in attempting to address this question are formidable, mainly because the ³He superfluid transition temperature (1 mK) is over 1,000 times lower than for ⁴He, but also because the maximum superflow rates should be orders of magnitude lower. In addition, there are sound theoretical arguments for supposing that, though undoubtedly superfluid in bulk, liquid ³He need not retain its superfluidity when in the form of a thin film. This is because of a characteristic length scale, the 'coherence' length ξ_0 , that represents the minimum distance over which the superfluid wavefunction can be significantly changed. For superfluid ⁴He, ξ_0 is comparable with the interatomic spacing so that superfluidity is to be expected (and is observed) down to extremely thin films

consisting of only a very few layers of atoms. For superfluid ³He, on the other hand, ξ_0 is of the same order as the thickness calculated for a 'saturated' film, (a film that is in equilibrium with bulk liquid below). Thus, the superfluidity of ³He films need not exist at all.

Positive evidence of superflow in ³He films, however, was reported in 1985 by

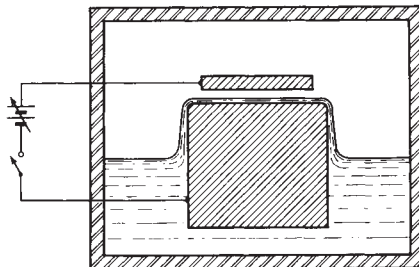


Fig. 1 Schematic diagram of the apparatus used for generation and detection of superfluid ³He film flow. When the potential difference across the capacitor plates is increased, liquid is drawn into the narrow space between the plates. The rate of flow is determined by measuring the change in capacitance. (From Davis *et al. Phys. Rev. Lett.* 60, 302–304; 1988).

the late J. G. Daunt (co-author of the historic paper cited above) and co-workers (*Phys. Rev. Lett.* 55, 1602–1605; 1985). Their experiment, based on measurements of the rate of filling or emptying of a small stainless steel beaker, caused consternation, because the superflow appeared to persist to temperatures above that of the bulk superfluid transition T_c^B . It had been anticipated, on the contrary, that 'finite-size effects' on scales comparable to ξ_0 would depress the superfluid transition temperature for the film T_c^F . More recent experiments by the same group (*Jap. J. appl. Phys.* 26, 145–146; 1987) suggest that the effect could have arisen from unsuspected temperature gradients.

The new experiment by Davis *et al.* is based on a very different approach. Their apparatus (Fig. 1) consists in essence of a horizontal parallel-plate capacitor partially immersed in liquid ³He. The potential difference between the plates can be changed and the capacitance can be monitored

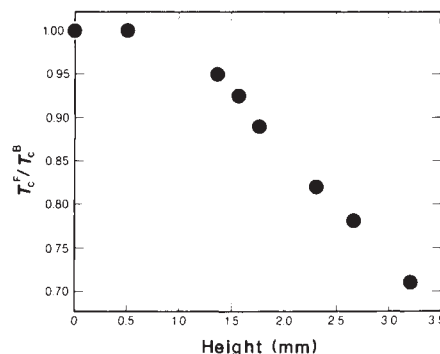


Fig. 2 Ratio of the superfluid transition temperatures for films and for bulk liquid ³He, T_c^F/T_c^B , as a function of the height of the film above the bulk liquid surface. At small heights (thick films), the transition temperatures are equal; for the thinner films at larger heights, T_c^F falls well below T_c^B , indicating the presence of finite-size effects. (From Davis *et al. Phys. Rev. Lett.* 60, 302–304; 1988.).

continuously. At the temperatures in question (less than 1 mK), the saturated vapour pressure is negligible, so that the space above the liquid contains a vacuum. When a potential difference V is applied between the plates, there is a tendency for the dielectric liquid helium to be drawn into the capacitor (a classical electrostatic phenomenon). In practice, this process would take an exceedingly long time at temperatures above T_c^B because the liquid is viscous. Below T_c^B , on the other hand, because of the superfluidity, progress towards a new equilibrium film thickness happens relatively quickly. The thickness of the film can be determined, to the remarkable accuracy of 0.1 nm within 1 s of observation time, by measuring the total capacitance, which is increased slightly by the liquid between the plates.

The ratio T_c^F/T_c^B is shown in Fig. 2 as a function of the height of the capacitor above the main liquid bath — in effect, as a function of the film thickness, which decreases with increasing height. It is immediately obvious that the superfluid transition temperature, at which liquid first flows into the capacitor in response to an increase of V , is equal to the bulk value for thick films, but falls by up to 30 per cent for thinner films, presumably because of the anticipated finite-size effects. Davis *et al.* also measured the critical (maximum) flow rates for the film; but these cannot be converted into the corresponding critical velocities because the lower capacitor plate is microscopically quite rough.

It seems that several interesting variations on the new technique are possible. Given its great sensitivity — the authors estimate it is better by a factor of 1,000 than earlier methods — it is reasonable to hope for rapid progress towards a detailed understanding of the superfluidity of the ³He creeping film. □

P. V. E. McClintock is in the Department of Physics, University of Lancaster, Lancaster LA1 4YB, UK.

Information processing

Neural populations revealed

Terrence J. Sejnowski

How are actions represented in the brain? Microstimulation of a discrete point in the deeper layers of the monkey superior colliculus produces a very fast saccade of the eyes with a particular direction and amplitude that depends only on the location of the microelectrode^{1,2}. The latency is short (20–30 ms) and the response is machine-like. This is reminiscent of an invertebrate 'command' neuron whose firing can cause a complex coordinated behaviour to occur (see page 278 of ref. 3). The pattern of activity in the superior colliculus preceding a saccade is, however, quite widespread⁴ and it has been generally assumed that the summation of the responses in the active population determines the eye movement. This hypothesis predicts that the neurons with peak activity in the motor map code the location to which the eyes move. Lee, Rohrer and Sparks, in their elegant experiment published on page 357 of this issue⁵, provide conclusive proof to the contrary, that the information coding for eye movement in the superior colliculus is represented by the distributed firing pattern of all the neurons that are activated. Such distributed representations of sensory information and motor commands may be quite common in the brain.

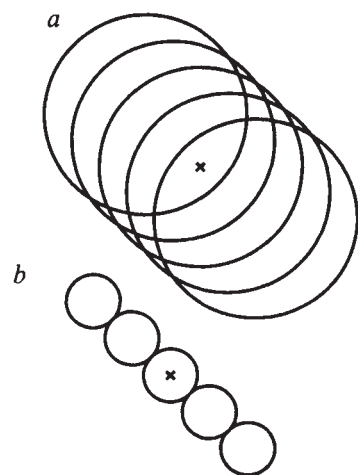
A region in the deeper layers of the superior colliculus about 2 mm in diameter containing about 100,000 neurons is activated just before a typical saccade. Imagine that each of these neurons has attached to it an arrow representing the direction and magnitude of the eye movement that occurs when the stimulating electrode is centred on that neuron. Imagine further that the 100,000 vectors attached to these neurons are added vectorially, each arrow weighted by the activity level of its neuron. The resultant vector accurately predicts how the eyes will move. Lee *et al.* show⁵ that when a subset of these neurons is deactivated by lidocaine, and can no longer contribute to the vector average, the eye movement will miss the target precisely as predicted by the resultant vector. This suggests a very large population of neurons contributes to the coding of a simple eye movement.

The use of such a large population to represent a single vector might seem wasteful. It can be shown, however, that a large population of neurons with broadly tuned response properties can code information more efficiently than a 'command' neuron representation in which only a single neuron is active for each eye movement^{6,7} (see figure). Moreover, accuracy is not lost in a distributed

representation, as long as only one eye movement is being represented at any given time. Thus, far from being a 'sloppy' way to organize information, coarse population coding is efficient, accurate and relatively insensitive to noise in the responses of single neurons.

Will this lesson learned in the oculomotor system generalize to other action systems, such as the control of the limbs,

Comparison of coarse and fine strategies for coding eye movements in the superior colliculus. Each point in the two-dimensional motor map in the intermediate layers of the superior colliculus represents a vector coding a particular magnitude and direction for an eye movement. Each circle encloses those neurons which contribute to a single eye movement, called the motor point image. The cross represents the centre of activation for a particular saccade. *a*, When the coding is coarse, many neurons contribute to a single eye movement and the same neurons contribute to many different eye movements (overlapping circles). *b*, For fine coding of an eye movement fewer cells are active and different eye movements are coded by different subpopulations (non-overlapping circles). The experimental evidence favours coarse coding in the superior colliculus⁵. Although more neurons are active in the coarse array, the movement is represented with greater accuracy than in a fine array with the same number of neurons. In general, the number of neurons required to code position to the same accuracy in a two-dimensional array decreases as $1/d$, where d is the size of the coding region⁶. A problem with coarse coding arises when two positions must be represented at the same time. The vector addition model predicts an eye movement which is halfway between the two positions.



which have more inertia and more degrees of freedom than eyes? Single neurons in the motor cortex of monkeys respond during reaching with broad tuning for direction of movement⁸. When the same vector-averaging procedure used in the superior colliculus is performed on a large population of cortical neurons, the direction of the population vector is congruent with the direction of movement of the hand. But the critical inactivation experiment has yet to be performed in motor cortex, partly because the organization of hand movements in the cortex is not as regular as the motor map of eye movements in the superior colliculus.

Distributed representations pose several problems for organizing complex systems like brains. How are decisions to be made when the relevant information is spread out over such a large population? How are so many neurons coordinated when precise timing is of the essence, as in playing a piano? How can new actions be learned without interfering with other actions produced by the same population of neurons? Experimental answers to most of these questions are not yet available. One source of intuition into the

properties of distributed representations that is available comes from network models composed of simple neuron-like processing elements. The design of distributed network models of the sensory-motor interface is just beginning⁹, but it is already clear that the issue of learning is central in developing and maintaining precise registration between sensory and motor representations (refs 10–12; D. Rumelhart, personal communication).

Network models of distributed sensory transformations are also being actively pursued, including models for the spatial transformation between retinal and head coordinates in parietal cortex¹³ and the representation of three-dimensional

shape in visual cortex¹⁴. Many counter-intuitive properties of distributed representations can now be demonstrated with computer simulations. Some of the mysteries of population codes may be revealed by a combined effort to both model and measure distributed sensory and motor representations in the brain. □

1. Robinson, D.A. *Vision Res.* **12**, 1795–1808 (1972).
2. Schiller, P.H. & Stryker, M. *J. Neurophysiol.* **35**, 915–924 (1972).
3. Bullock, T.H., Orkand, R. & Grinnell, A. *Introduction to Nervous Systems* (Freeman, San Francisco, 1977).
4. Wurtz, R.H. & Goldberg, M.E. *J. Neurophysiol.* **35**, 575–586 (1972).
5. Lee, C., Rohrer, W.H. & Sparks, D.L. *Nature* **332**, 357–360 (1988).
6. Hinton, G.E., McClelland, J.L. & Rumelhart, D.E. in *Parallel Distributed Processing. Explorations in the Microstructure of Cognition* (eds Rumelhart, D.E. & McClelland, J.L.) 77–109 (MIT, Cambridge, 1986).
7. Heiligenberg, W.J. *Comp. Physiol.* **A161**, 621–631 (1987).
8. Georgopoulos, A.P. *et al. Science* **233**, 1416–1419 (1986).
9. Pellionisz, A. & Llinas, R. *Neuroscience* **7**, 2949 (1982).
10. Kuperstein, M. *Science* **239**, 1308–1311 (1988).
11. Reeke, G.N. Jr & Edelman, G.M. *Daedalus* **143** (1988).
12. Knudsen, E.I. *J. Neurosci.* **5**, 3094–3109 (1985).
13. Zipser, D. & Andersen, R.A. *Nature* **331**, 679–684 (1988).
14. Leaky, S.R. & Sejnowski, T.J. *Soc. Neurosci. Abstr.* **13**, 1451 (1987).

Terrence J. Sejnowski is a professor of Biophysics, Biology, Electrical and Computer Engineering, and Computer Science at the Johns Hopkins University, Baltimore, Maryland 21218, USA.

Biochemical cycles

Natural volatilization of tin

Peter Craig

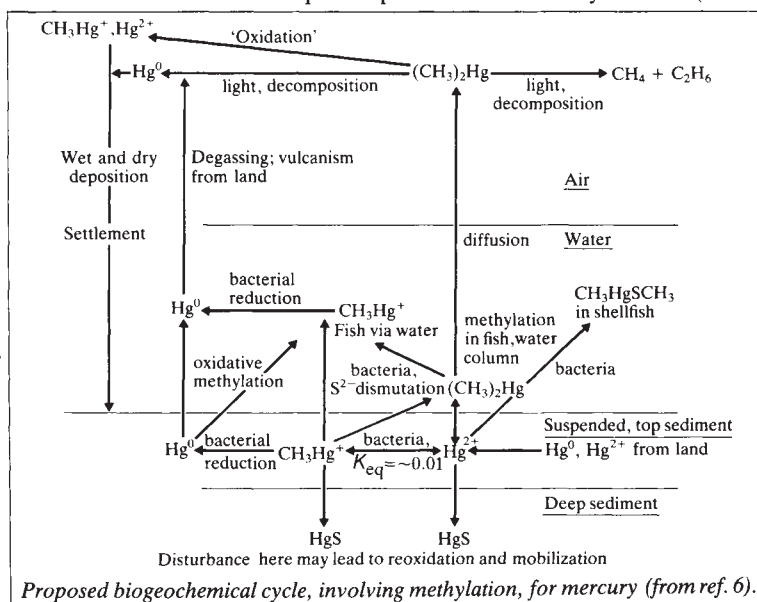
It is well known that there are natural biochemical cycles for carbon and nitrogen, and recent work has shown that there is analogous biogeochemical cycling of some metallic and metalloidal elements¹. Best known among these are the cycles for arsenic and mercury (see figure), involving substantial chemical transformation of these elements to volatile methylated forms which diffuse to the atmosphere. Donard and Weber's work, reported on page 339 of this issue², significantly extends our knowledge about metallic cycles in the biosphere in revealing a new pathway for volatilization for one metal, tin, from aqueous to atmospheric phases. This pathway, which does not involve methylation, may be of significance for other metals.

In Donard and Weber's experiment, inorganic tin compounds (SnCl_4 and tin metal) incubated for 6 months in an anoxic medium containing decaying algal material, were slowly converted to hydrophobic volatile stannane (SnH_4). This stannane then diffuses through the water above, without being oxidized, into the head space of the environmental chamber used in the experiment, in a way that could plausibly occur in real environments. One source of tin for such a process in the natural environment is the tributyltin antifouling paint common in harbours and marinas.

Interest in metal methylation (biomethylation) first arose because of pollution incidents in the late 1950s. For example, reductions in Canadian and Swedish bird populations occurred because of poor hatching rates due to eggshell thinning induced by anthropogenic methyl mercury from seed dressings³. Fish in various locations were found to have parts per million levels of methyl mercury, despite having had no apparent contact with this compound⁴. It was soon demonstrated that inorganic mercury added to aquarian sediments is partially but quickly methylated to methyl mercury⁵. Although the pollution levels in sediments are very low, food-chain effects that increase the concentration of the stable methyl mercury compounds led public-health authorities to impose maximum permissible levels of mercury in fish for human consumption.

More recently, interest has moved from pollution effects to the role of

methylation in assisting the bulk transport of metals in the environment. For mercury, methylation to long-lived, non-volatile CH_3Hg^+ can be followed by dismutation to volatile, diffusible hydrophobic $(\text{CH}_3)_2\text{Hg}$ by interaction with sulphides in sediments. It has been estimated that up to 12 per cent of the methyl



mercury per year could be removed from a sediment and diffused through the water layer to the atmosphere⁶. No natural alkylation by groups other than methyl has yet been observed. Metal hydride formation has only recently been discovered.

Increasingly sophisticated analytical techniques are being developed to identify the methyl metals. Gas chromatography, mass spectroscopy, atomic absorption spectroscopy and other techniques can analyse methyl metals at concentrations of a few parts per billion (10^9) or below in various environmental matrices. Naturally methylated elements thus identified include mercury, tin, arsenic, antimony, germanium, thallium, selenium, tellurium and, less clearly, lead and silicon⁷. Much work has been carried out on lead, but the data are not easy to interpret, and biomethylation is hard to recognize, because of general environmental contamination by methyl-lead gasoline additives.

The latest twist in the ability of the environment to synthesize unusual inorganic compounds is revealed by the work of Donard and Weber in this issue², showing that certain algae can convert inorganic tin compounds to stannane. This is the first demonstration of the environmental production of a simple

metal hydride, and several issues are raised by this observation. As tin is a Group IV element, as is carbon, it could be that the production of stannane arises by the substitution of tin for carbon at critical points in the pathway of methane biosynthesis. It is significant that methanogenic activity was also observed in this experiment, as this result suggests that methanogenic organisms are responsible for stannane formation. That some methyl tins are also formed adds to this suspicion.

The existence of these apparently unstable compounds under natural conditions (normally they oxidize quickly in air) could be explained by the fact that they originate in anoxic environmental compartments, and also because their low concentration reduces the rate of intermolecular collisions leading to oxidation. Donard and Weber's work also appears to be the first demonstration of the natural formation of a simple heavy metal hydride, although the mixed metalloidal species $(\text{CH}_3)_2\text{AsH}$ can be formed microbially from inorganic arsenic compounds. This has important implications for tin transport in the environment.

Naturally occurring mixed organometallic hydrides have been identified⁸ in Baltimore harbour. These compounds, $(\text{CH}_3)_n\text{SnH}_{4-n}$ ($n = 1-3$) and $(\text{CH}_3)_3\text{Sn}$, could have biogenic or anthropogenic origins. The highest levels found for the methyl stannanes are up to about $0.5 \mu\text{g l}^{-1}$ for each species, well within the limits of detection by modern equipment. Up to about $1.0 \mu\text{g l}^{-1}$ $(\text{CH}_3)_3\text{Sn}$ is also found. Laboratory experiments confirm⁸ that inorganic tin can be methylated and reduced to a methyl tin hydride by an aerobic strain of *Pseudomonas* 244 which was isolated from Chesapeake Bay, and small quantities of $(\text{CH}_3)_3\text{Sn}$ and $(\text{CH}_3)_3\text{SnH}$ were also detected in the incubation. It will be interesting to see if there are other simple metal hydrides that occur naturally. □

1. Nriagu, J.O. & Donaldson, C.I. (eds) *Toxic metals in the Atmosphere* (Wiley, New York, 1986).
2. Donard, O.X.F. & Weber, J.H. *Nature* **332**, 339-341 (1988).
3. Fimreite, N. in *The Biogeochemistry of Mercury in the Environment* (ed. Nriagu, J.O.) 601-27 (Elsevier, Amsterdam, 1979).
4. Rudd, J.W.M., Furutani, A. & Turner, M.A. *Appl. Environ. Microbiol.* **40**, 777-82 (1980).
5. Jensen, S. & Jernelov, A. *Nature* **223**, 753-754 (1969).
6. Craig, P.J. (ed.) *Organometallic Compounds in the Environment* (Longman, London, 1986).
7. Craig, P.J. & Moreton, P.A. *Mar. Pollut. Bull.* **15**, 406-408 (1984).
8. Jackson, J.-A.A. et al. *Environ. Sci. Technol.* **16**, 110-119 (1982).

Peter Craig is in the School of Chemistry at Leicester Polytechnic, PO Box 143, Leicester LE1 9BH, UK.

A.R.J.P. Ubbelohde (1907–1988)

A.R.J.P. UBBELOHDE, who died on 7 January, was an exceptionally versatile scientist and a man of broad cultural tastes. A pig farmer, a keen fisherman, an enthusiastic swimmer, a collector of antiques, a connoisseur of wine and a classical scholar of no mean distinction: he was all these and more, but his scientific reputation rests on his extensive contributions, principally experimental, to a wide range of physico-chemical topics encompassing the behaviour of solids, liquids and gases. Only a small circle knew him as Paul Ubbelohde. To many he was "Ubbel-bubble", a description he silently approved as it recognized his perpetual effervescence, boundless enthusiasm and enormous intellectual energy.

He was born on 14 December 1907, and was educated at St Paul's School and Christ Church, Oxford. In the early 1930s, he spent a year at Göttingen where he saw at first hand the impressive physico-chemical legacy and continuing efforts of the great German physical scientists including Nernst, Planck, Eucken and Clusius. But, as he often declared, the most formative period of his scientific life was spent at the Royal Institution (1935–40), where William Bragg allowed him and a galaxy of gifted contemporaries (including K. Lonsdale, J.M. Robertson and H.A. Jahn) to roam freely into whichever field took their fancy — a far cry from the present-day ethos of target-orientated, interdisciplinary research centres.

His publications at that time reveal his unusual range of interests, most of which he pursued later: hydrogen uptake by metals, which led to refinements of the phase rule; melting and crystal structure, the subject of one of his many books; the ferroelectric behaviour of Rochelle salt; the influence of isotopic substitution upon crystalline properties; and occurrence of induction periods in the combustion of hydrocarbons. Ubbelohde, although a junior among his associates — he referred to himself as "the Benjamin of the family" — was capable of mobilizing the interest, skills and energy of all around him. His first book, *Modern Thermodynamics*, was published during this period.

At Queen's University, Belfast (1945–54) he built up a vibrant department of physical chemistry. His own research included transport properties and the combustion of gaseous hydrocarbons; the diffusion, ultrasonic dispersion and viscosity of liquids; and the interaction of alkali metals with aromatic hydrocarbons. He also studied the intercalation of guest species by graphitic hosts, a phenomenon discovered in West Germany in the early nineteenth century, and resurrected by U. Hoffman and others at Heidelberg immediately after the Second World War. Ubbelohde's contributions to graphite

and the crystalline solids it forms with many guests are of lasting value. Apart from recognizing the enormous range of electrical and magnetic properties exhibited by these intercalates, he also stimulated others to prepare synthetic analogues of graphite, like stress-annealed pyrographite, which opened a new chapter in the design of layered solids with advantageous thermal and mechanical



properties. In 1954 he was appointed professor of thermodynamics and, from 1961 until his retirement in 1975, head of the department of chemical engineering and chemical technology at Imperial College, London.

Ubbelohde's presence at a scientific meeting always made it interesting. He spoke in clarion, authoritative terms often with humour and generally with reference to arcane events usually of Greek or Roman provenance. There were, however, occasions when he lapsed into obscurantism for which he could be forgiven. But those who listened to him always felt that they were the better upon hearing his views.

He knew how to capture an audience both in the spoken and written word. His books contain evocative passages. Thus in introducing *Melting and Crystal Structure*, he remarks: "Traditionally the interest of the early Directors of the Royal Institution in the effect of pressure on melting stems from Humphry Davy's prowess as a skater". And in *Man and Energy*, he illustrates the second law of thermodynamics with the quip that "An untidy nursery at the end of the day will not tidy itself up spontaneously", and the moral dilemma of the scientist with the recollection that "It is notable that Leonardo was so concerned about possible abuses of his invention of a submarine by tyrants that he took pains not to publish it". Ubbelohde was a natural stylist. He did and said things in a memorable way.

J.M. Thomas

J.M. Thomas is Director of the Royal Institution, 21 Albemarle Street, London W1X 4BS, UK.

Daedalus

Abstract concrete

CONCRETE, says Daedalus, is an embryonic fractal material. Its cheap, coarse gravel particles pack tightly together; their interstices are filled with almost equally cheap sand; the expensive binding cement is needed only to fill the residual space between the sand grains. So Daedalus is taking this logic to its ultimate conclusion. He is devising a graded range of particles from the coarsest gravel to the finest imaginable dust. Each grade of particle will pack neatly in the interstices of the grade above, while providing in its own interstices a tight fit for the particles of the grade below. At each stage of packing the unfilled volume diminishes, and in the ultimate mathematical limit it approaches an unimaginably tortuous fractal web of infinite surface area but zero volume. The final concrete, therefore, contains 100 per cent of cheap aggregate and 0 per cent of expensive binder.

In fact, of course, one is prevented from pursuing this reticulated logic to its final nightmarish conclusion by atomic dimensions and by the limits of fine-grinding. About seven or eight grades of nesting particle sizes span the feasible technological range, and their packing poses some intriguing statistical problems. But the final fractal lattice will have only a few parts per million of free volume between its close-packed tiers of nesting particles. So little binder will be needed that expensive, high-tenacity types like epoxy, cyanoacrylate or even polyimide will become economic.

Fractal concrete cannot be made in the traditional way, by adding the liquid to the solid and stirring. Viscous binding and drag between the closely packed particles would make the mixture far too stiff to stir. Instead the solids must be mixed dry and compacted in place. The calculated small volume of binder should then be poured on. The enormous capillary attraction of the atomically close particles will suck it in and distribute it evenly around the tortuous but tiny free volume between them. It will soon set, giving a concrete of unprecedented strength and resistance.

Many other composite materials could be improved by a similar approach. A fractal fibreglass, for example, in which small fibres nest in the hollows between big ones, and smaller ones still in the hollows between them, should be much stronger than any monodisperse array. Daedalus is even devising an ultra-healthy fractal muesli, whose carefully sized grains of nut, fragments of cereal and fine vegetable fibres pack so densely in the bowl that they can be fully wetted with the absolute minimum volume of saturated-fatty milk.

David Jones

Microwave syntheses for superconducting ceramics

SIR—The extraordinary research efforts that have followed the discovery of the high-critical-temperature superconducting ceramics based on copper oxides^{1,2} have concentrated primarily on the unique physical properties of these materials, their structural characterization and the development of theoretical models³. The development of alternative procedures for synthesizing these compounds has received less attention.

Recently we have discovered that several inorganic oxides including CuO, absorb microwave radiation of 2,450 MHz, which is the frequency used in domestic microwave ovens, very strongly⁴. As a result of this strong absorption, 1–5-g samples of CuO achieve temperatures in excess of 550 °C after a 1-min exposure in a commercially available microwave oven operating at a power level of 500 W. We have used this oven to synthesize mixed metal oxides, including those with superconducting properties.

The procedures adopted and the times required for the experiments can be gauged by giving details of the specific synthesis of La₂CuO₄. La₂O₃ (12.28 g) was mixed with CuO (3.00 g) by mechanical shaking and intimately mixed using an agate mortar and pestle. The mixture was placed in an alumina crucible and supported on a firebrick in the oven. A minute after the oven was turned on the mixture was glowing orange; this glow was maintained for a further 9 min and the mixture became molten, whereupon the oven was switched off. The product was allowed to cool and then ground to a fine powder, which was shown to be mainly La₂CuO₄ by X-ray powder diffraction. To convert the mixture into pure La₂CuO₄ it was heated for a further 30 min using variable powers of 130–500 W and then cooled and ground. The refined cell constants of the La₂CuO₄ were: $a = 5.354(1)$ Å, $b = 5.402(1)$ Å and $c = 13.149(1)$ Å (ref. 5).

Conventional heating techniques would typically require 12–24 h for this synthesis, so synthesis in a microwave oven leads to the isolation of a pure product in a fraction of the time usually required for conductive heating methods. Furthermore, X-ray diffraction experiments indicate the final product is not only of high purity but also of high crystallinity.

La_{1.85}Sr_{0.15}CuO₄ was similarly synthesized using CuO, La₂O₃ and SrCO₃ in the appropriate stoichiometric ratios. The first microwave heating cycle was completed in 6 min, and further cycles of 23 and 5 min lead to a pure sample, as determined by X-ray diffraction experiments with a tetragonal cell of dimensions: $a = b = 3.7751(4)$ and $c = 13.224(3)$ Å (ref. 6).

YBa₂Cu₃O_{7-x} was synthesized from

CuO, Y₂O₃ and Ba(NO₃)₂ in the appropriate stoichiometric ratios. The mixture was heated in a modified microwave oven which enabled the liberated NO₂ gas to be safely expelled. When treated with microwave radiation at a power level of 500 W for 5 min all the NO₂ was liberated and a powder diffraction pattern showed that no Ba(NO₃)₂ remained in the sample. The mixture was reground and exposed to microwaves (130–500 W) for a further 15 min, reground and exposed for an additional 25 min. Powder diffraction studies showed that the major component of the mixture at this stage was YBa₂Cu₃O_{7-x}, but there were some lines of low intensity due to Y₂BaCuO₅ (ref. 7). An additional exposure to microwaves for 25 min resulted in an essentially pure sample, with a tetragonal unit cell of dimensions $a = b = 3.861(2)$, $c = 11.789(7)$ Å (ref. 7). The conversion of

this tetragonal form into the superconducting orthorhombic form is achieved by slow cooling in the conventional manner⁷.

Having established that microwave heating provides a convenient and rapid mode of synthesis for the new class of high-temperature ceramic superconducting materials, we are now investigating the physical properties of these materials to establish whether this unusual mode of synthesis leads to any unusual properties.

D.R. BAGHURST
A.M. CHIPPINDALE
D.M.P. MINGOS

*Inorganic Chemistry Laboratory,
University of Oxford,
South Parks Road, Oxford OX1 3QR, UK*

1. Bednorz, J.G. & Müller, K.A. *Z. Phys.* **B64**, 189–193 (1986).
2. Wu, M.K. *et al.* *Phys. Rev. Lett.* **58**, 908–911 (1987).
3. Garwin, L. & Cambell, P. *Nature* **330**, 611–614 (1987).
4. Baghurst, D.R. & Mingos, D.M.P. *J. chem. Soc. chem. Commun.* (submitted).
5. Longo, J.M. & Raccach, P.M. *J. solid state Chem.* **6**, 526 (1973).
6. Bednorz, J.G. *et al.* *Mater. Res. Bull.* **22**, 819 (1987).
7. Kini, A.M. *et al.* *Inorg. Chem.* **26**, 1836 (1987).

Punctuation in perspective

SIR—The important issue raised by Eldredge and Gould¹ in response to my News and Views comment² on Sheldon's paper³ concerns the role of 'species selection'. I wrote "I do not think that the attempt to explain morphological evolution by species selection can survive . . . But there never was much sense in the idea anyway". In their reply, Eldredge and Gould agree that it makes no sense, but say that this characterization is entirely of my own making. Not so.

To show this, I must quote. Gould wrote⁴ "Speciation is not always an extension of gradual allelic substitution to greater effect, but may represent, as Goldschmidt argued, a different style of genetic change — rapid reorganisation of the genome, perhaps non-adaptive. Macroevolutionary trends do not arise from the gradual adaptive transformation of populations, but usually from a higher-order selection operating on groups of species, while the individual species themselves generally do not change". Expressing a similar idea, Stanley wrote⁵ "If rapidly divergent speciation interposes discontinuities between rather stable entities (lineages), and if there is a strong random element in the origin of these discontinuities (in speciation), then phylogenetic trends are essentially *decoupled* (his italics) from phyletic trends within lineages . . . Then what determines the course of most phylogenetic trends must be a selection process in which species are the units". These "trends" must be morphological, because morphology is all, or almost all, that we see in the fossil record. I am delighted that Eldredge and Gould have abandoned species selection as a significant cause of morphological evolution,

but they really must not pretend they never said otherwise.

I say that species selection is the important issue, because big claims have been made for punctuational theory. The article by Gould quoted above is entitled "Is a new and general theory of evolution emerging?" (the answer of course is yes), and the opening words are "The modern synthesis, as an exclusive proposition, has broken down". Such a claim cannot be justified by what Eldredge and Gould say in reply to my article. The insistence that allopatric speciation is "the very heart and soul of punctuated equilibria", or that the rates of evolution in lineages leading to new major taxa are orders of magnitude greater than those observed (for example) in trilobites, imply no break with the modern synthesis. The former concept was Mayr's⁶ central thesis, and the latter was Simpson's⁷: Mayr and Simpson, of course, were two of the major architects of the synthesis. The only new idea in the theory of punctuational equilibria that could, if correct, justify claims of a new evolutionary paradigm is the idea of species selection as outlined in the quotations above.

Two lesser points. Eldredge and Gould suggest that the changes described by Sheldon are too small to shed much light on the origin of taxa above the species level. This may, or may not, be true, but it is an odd claim for Eldredge, at least, to make. In his book⁸ explaining the theory of punctuated equilibria, the only example discussed at any length is his own study of changes in the number of lenses in the eyes of trilobites. Why a change from 18 to 17 columns of lenses in the eye is relevant, whereas a change from 11 to 13 pygidial ribs³ is not, defeats me. (It was, in fact, the close similarity between the traits

studied by Eldredge and by Sheldon that led me to concentrate, in my article, on Sheldon's data³ and to ignore other equally relevant data, such as Gingerich's⁹.)

The second point is their claim that all parties to the debate "have focussed on the statistics of measured morphological change, and not published compendia of names". Would that it were so, but it is not. Stanley, in the best-documented case for punctuation⁵, relies almost entirely on the duration in the fossil record of named taxa — sometimes species, but often genera or families.

I think there is much of importance in the writings of the punctuationists. Stasis is a real phenomenon, and a fascinating one. The suggestion that different rates of speciation and extinction — depending, perhaps, on differences in breeding systems, dispersal¹⁰ or degree of ecological specialization¹¹ — have influenced patterns of diversity must surely be true, and it will be interesting to work out more of the details. But all this can sit very happily within the current neo-darwinist view. It seems, from their response to my article, that Eldredge and Gould now recognize this. Welcome back!

JOHN MAYNARD SMITH

*School of Biological Sciences,
University of Sussex,*

Falmer, Brighton, Sussex BN1 9QG, UK

1. Eldredge, N. & Gould, S.J. *Nature* **332**, 211–212 (1988).
2. Maynard Smith, J. *Nature* **330**, 516 (1987).
3. Sheldon, P.R. *Nature* **330**, 561–563 (1987).
4. Gould, S.J. *Palaeobiology* **6**, 119–130 (1980).
5. Stanley, S.M. *Macroevolution* (Freeman, 1979).
6. Mayr, E. *Systematics and the Origin of Species* (Columbia University Press, 1942).
7. Simpson, G.G. *Tempo and Mode in Evolution* (Columbia University Press, 1944).
8. Eldredge, N. *Time Frames* (Heinemann, London, 1986).
9. Gingerich, P.D. *A. J. Sci.* **276**, 1–28 (1976).
10. Jablonski, D. & Lutz, R.A. *Biol. Rev.* **59**, 21–89 (1983).
11. Vrba, E.S. in *Living Fossils* (eds Eldredge, N. & Stanley, S.M.) (Springer, Berlin, 1984).

AIDS incubation period in haemophiliacs

SIR—Ekert's assertion¹, apparently accepted by Turner *et al.*², that the rate of conversion to AIDS of haemophiliacs infected with the human immunodeficiency virus (HIV) is a tenth of the other risk groups is unlikely to be correct.

Most haemophiliacs in the United Kingdom have now been tested and 1,087 were listed as positive in England, Wales, Scotland and Northern Ireland in December 1987. Taking into account the few positives who have so far been missed, the actual total of UK haemophilic positives is probably between 1,100 and 1,200. By December 1987 a cumulative total of 70 UK haemophiliacs had been reported with AIDS. So far, therefore, about 6 per cent of haemophilic infectives have progressed to AIDS.

By December 1987 a total of 1,227 AIDS cases had been reported in the United Kingdom. If 50,000 people are infected this suggests an overall AIDS conversion rate of 2.45 per cent; if the real

Solitons and energy transfer in DNA

SIR—There has been discussion recently in this journal¹ on the possibility that solitons (solitary waves) play a role in the properties of DNA. The observations that stimulated this discussion^{2,3}, have since been questioned^{4,5}. Most attention has been focused on the theoretical case for the propagation of solitons along DNA^{6,7} and very little has been given to the conditions necessary for their genesis, for which a non-linear local disturbance of the equilibrium condition above a threshold value is required. Bednar⁸, in discussing the energy deposition process of ionizing radiation leading to radiation damage in biological systems, postulates the genesis of solitons.

This process, in which up to 100 eV can be deposited in segments of the DNA molecule a few base pairs (bp) long, results in coherent multiple excitations and ionizations⁹ and satisfies the requirement for a significant departure from thermal equilibrium. We consider here the possibility that solitons may provide a mechanism for transferring energy in DNA.

We have reviewed elsewhere¹⁰ experiments in which strand breaks are induced in DNA primarily as a result of the direct absorption of ionizing energy^{11–13} and suggest that such results require an explanation in terms of the transfer of large amounts of energy (bond-breaking energy) over larger distances in molecular terms (thousands of bp) than has previously been generally thought possible in DNA radiation chemistry.

We propose that solitons generated by the energy deposition process could pro-

vide a mechanism for this transfer. We suggest that vibrational modes characteristic of DNA (for example sugar puckering¹⁴) couple to excitation energy or charge to enable energy to 'ride' along the molecule.

In other words the soliton simultaneously creates both the environment necessary to support excitation energy or charge, and the conditions for its translation. This proposal is reminiscent of the solvated electron concept in which the electron couples with adjacent water molecules to form a comparatively stable 'polaron'. The effectively one-dimensional nature of DNA permits a translational wave to propagate, so transferring the energy either as an exciton (coupled electron-hole pair) or as a charge (polaron).

Energy localization, and hence damage, may result from the collapse of the soliton at some discontinuity in the structure of the DNA molecule, such as a B/Z transition, an intercalated or otherwise bound molecule (particularly with energy or charge accepting character) or as the result of the coincidence of two passing solitons.

Although there is no single piece of evidence in radiobiology that explicitly demands an explanation of the kind we have given, the results mentioned above in which the DNA is essentially native^{11–13} require explanation, and this it seems to us must lie outside the present conventional radiobiological approach in which damage to the DNA is assumed to occur at the sites of initial energy deposition.

The relevance of studies on isolated DNA to radiobiology is indicated by measurements of single and double-strand breaks in plasmid DNA irradiated in 'dry' films. The energy requirements for such breaks are the same as for SV40 DNA (which is similar in size to the plasmids) irradiated in cells¹⁵. This we interpret to imply a common mechanism between the two systems, namely that of the direct absorption of ionizing energy which dominates in the dry films.

There are profound implications for radiobiology if mechanisms involving solitons are important in irradiated DNA. The possibility that the distribution of damage in native DNA is mediated at least in part by the microstructure of the DNA double helix, and is not solely a function of the spatial distribution of the initial ionizing events, is one that has received scant attention to date, despite having been proposed in 1972 by Neary¹¹.

Solitons have been mentioned in attempts to explain the basic functions of DNA in cells, for example protein production and conformational transitions. How the energy needed for the processes and control of DNA replication and synthesis is transferred is obscure, as, indeed, in

total of infectives is 100,000, as I believe³, the overall rate would be 1.2 per cent.

Because haemophiliacs include a much higher proportion of the very young and very old than other at-risk groups, overall they will progress to AIDS more rapidly. As UK haemophiliacs received blood products from the United States they may, in general, have been infected slightly earlier than other groups in the United Kingdom.

Whereas we know nearly all of the haemophilic HIV-positives, in the case of other at-risk groups the known HIV-positives are a highly biased sample, because many already show signs of serious illness.

MALCOLM REES

*John Hunter Clinic,
St. Stephens Hospital, Fulham Road,
Chelsea, London SW10 9TH, UK*

1. Ekert, H. *Nature* **329**, 494 (1987).
2. Turner, M.J. *et al.* *Nature* **330**, 702 (1987).
3. Rees, M. *Nature* **326**, 343–345 (1987).

radiobiology, is the relationship between energy deposition, molecular damage to cellular components, and biological 'end-point'.

K.F. BAVERSTOCK
R.B. CUNDALL

MRC Radiobiology Unit,
Chilton, Oxon, UK

1. Maddox, J. *Nature* **324**, 11 (1987).
2. van Zandt, L.L. *Phys. Rev. Lett.* **57**, 2085 (1986).
3. Englander, S.W., Kallenbach, N.R., Heeger, A.J., Krumhansl, J.A. & Litwin, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7222-7226 (1980).
4. Frank-Kamenetskii, M. *Nature* **328**, 108 (1987).
5. Gabriel, C. *et al.* *Nature* **328**, 145-146 (1987).
6. Davydov, A.S. *Phys. Script.* **20**, 387-394 (1979).
7. Davydov, A.S. & Zolotarev, A.V. *Phys. Script.* **28**, 249-256 (1983).
8. Bednar, J. *Int. J. radiat. Biol.* **48**, 147-166 (1985).
9. Klein, G. & Voltz, R. *Int. J. radiat. phys. Chem.* **7**, 155-174 (1975).
10. Baverstock, K.F. & Cundall, R.B. *Int. J. radiat. phys. Chem.* (In the press).
11. Neary, G.J., Horgan, V.J., Bance, D.A. & Stretch, A. *Int. J. radiat. Biol.* **22**, 525-537 (1972).
12. Baverstock, K.F. *Int. J. radiat. Biol.* **47**, 369-374 (1985).
13. Linz, U. & Stöcklin, G. *Radiat. Res.* **101**, 262-278 (1985).
14. Takeno, S. & Homma, S. *Prog. theor. Phys.* **77**, 548-562 (1987).
15. Baverstock, K.F. & Will, S. In *Radiation Research: Proc. 8th Intl Congr. Radiation Research, Edinburgh* (eds Fielden, E.M. *et al.*) 10 (1987).

Wood treatment used in Cremonese instruments

SIR—There has long been speculation about the methods and materials used by the seventeenth and eighteenth century schools of Italian violin makers of whom Antonio Stradivari of Cremona is the most celebrated. In particular, the wood treatment and varnish are vital to the outstanding appearance of his instruments and possibly also to their acoustic quality. Chipura¹ has argued persuasively that a wood treatment likely to have been used by the Cremonese craftsmen would have involved the application of Roman hydraulic cement, a material that would have been familiar to them. This cement is made by interacting lime with the locally available pozzolana volcanic ash² and imparts a hard and durable surface.

To test this theory experimentally, we used scanning electron microscopy (SEM) coupled with elemental analysis by energy-dispersive X-ray spectroscopy (EDAX)³. Examination of a sample of red

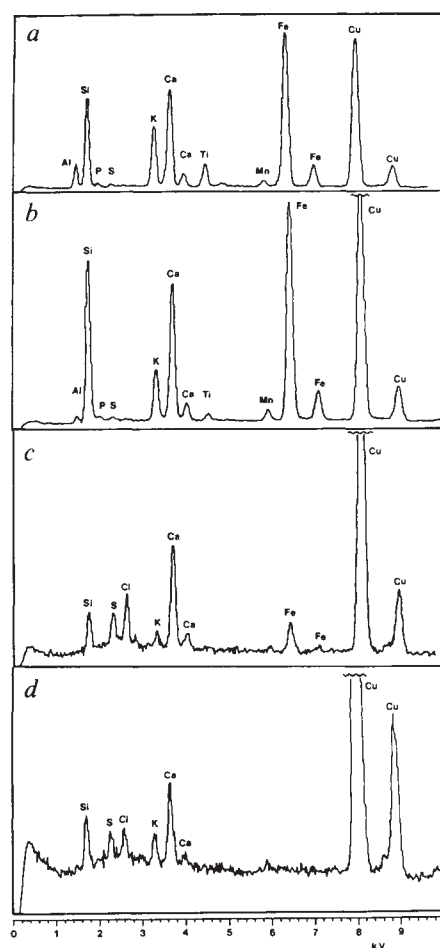


Fig. 1 EDAX spectra of typical particles from *a*, red pozzolana, *b*, Stradivari wood surface, *c*, varnish and *d*, wood interior from a Link Systems energy dispersive X-ray emission spectrometer attached to a JEOL JEM 200 CX transmission electron microscope. Accelerator voltage, 200 keV. Samples (held in a graphite single tilt holder) are tilted by 45° so their upper surfaces face the spectrometer. Spectra recorded over 100-s lifetime at a resolution of 20 eV per channel. Electron beam focused to spot sizes from 2 µm upwards to analyse single particles or groups of particles supported on thin holey carbon films on copper grids.

pozzolana from Valentana by EDAX showed the heterogeneity of the mineral; although the proportions of the elements varied, most particles contained Al, Si, P,

S, K, Ca, Ti, Mn and Fe; Fig. 1*a* shows a typical spectrum (the copper peaks result from the copper grid on which the particles were supported). A sample of yellow pozzolana from Bologna yielded similar spectra but with higher proportions of sulphur.

Generous contributors have provided fragments of authenticated Cremonese instruments. An SEM section (Fig. 2) of a rib from a 1711 Stradivari cello (from W.E. Hill and Sons) shows a dense particulate material (P) between the varnish layer (V) and the wood (W). A typical EDAX spectrum of the particulate material (Fig 1*b*) indicates the presence of Al, Si, P, S, K, Ca, Ti, Mn and Fe; there is a clear resemblance to the pozzolana spectrum in Fig. 1*a*. By comparison, particles dispersed in the varnish layer contained only Si, S, Cl, K, Ca and Fe (Fig. 1*c*), with a much lower proportion of iron than in pozzolana. Particles taken from the wood interior (Fig. 1*d*) contained only Si, S, Cl, K and Ca.

It has been suggested that this analysis could point equally to another material of volcanic origin, pumice, which was, and still is, used as a fine abrasive. There are two arguments against this interpretation. First, it seems improbable that this use of pumice would build up the observed particle thickness on the wood surface (Fig. 2). Second, an EDAX spectrum of a powdered pumice sample showed Al, Si, S, K, Ca, Ti and Fe, with much more silicon and much less iron than in Figs 1*a* or 1*b*.

Although these findings indicate that the range and proportions of elements present on the wood surface of this Stradivari instrument are compatible with the presence of pozzolana ash, Chipura's suggestion that the pozzolana was applied in the form of Roman cement cannot be verified; in such a case the surface particles should show a greatly enhanced calcium content. We are extending these investigations to other fragments from instruments of the period hoping not only to be able to define the nature of the wood surface treatment more precisely but also to reveal the methodology of the Cremonese craftsmen.

CLAIRE Y. BARLOW

Department of Materials Science
and Metallurgy, Pembroke Street,
Cambridge CB2 3QZ, UK

PETER P. EDWARDS*

G. ROBERT MILLWARD

RALPH A. RAPHAEL

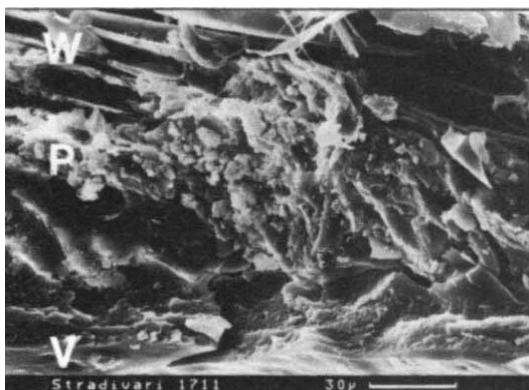
DAVID J. RUBIO

University Chemical Laboratory,
Lensfield Road,
Cambridge CB2 1EW, UK.

*To whom correspondence should be addressed.

1. Chipura, J. *Strad* **95**, 161 (1984).
2. Lea F.M. *Chemistry of Cement and Concrete* (Arnold, London, 1970).
3. Skoog, D.A. & West D.M. *Principles of Instrumental Analysis* 350 (Holt, New York, 1971).

Fig. 2 SEM showing part of a through-thickness fracture surface running normal to the surface of the instrument. V, varnish layer, the outside of the instrument, containing few obvious inclusions and 5-10 µm thick. P, particulate region, up to 100 µm thick with most particles under 4 µm in diameter, mainly 0.5-2 µm. Most are rough in shape, some very nearly spherical, and in general aspect ratios are around unity. There is some visible porosity, with cracks and voids up to 1 µm. W, (maple) wood penetrated by particles to a depth of up to 30 µm.



Ultra-high energy radiation from young supernovae

MOST cosmic rays are believed to originate in first-order Fermi acceleration at blast shocks in supernova remnants. The acceleration is fully efficient after about 1,000 years, and is restricted to $E_p < 10^{14}$ eV (ref. 1). There are suggestions²⁻⁵ that cosmic rays with energy up to 10^{18} eV might be accelerated in young supernova remnants for several years after the explosion.

A pulsar inside a supernova would release energy by magnetic dipole radiation or its equivalent at an approximate rate⁶

$$L = 4 \times 10^{43} B_{12}^2 R_{10}^6 P_{ms}^{-4} \times (1 + R_{10}^6 B_{12}^2 t / 16 P_{ms}^2 \text{ yr})^{-2} \text{ erg s}^{-1} \quad (1)$$

where P_{ms} , B_{12} and R_{10} are the initial period (in ms), the surface magnetic field (in 10^{12} G) and the radius (in 10 km) of the pulsar and t is the time since explosion. The time dependence assumes spindown by magnetic dipole losses only. A fraction of the pulsar luminosity might be used to accelerate protons deep inside the supernova shell. The maximum achievable proton energy in the pulsar wind shock model⁵ is $E_p^{\text{max}} \approx 1.9 \times 10^6 B_{12} P_{ms}^{-2}$ TeV. Charged particles with energy less than $E_{\text{esc}} \approx 10^7 B_{12} u_9^{1/2} / \xi P_{ms}^2$ TeV, where u_9 is the expansion velocity of the shell in units of 10^9 cm s^{-1} and $\xi > 1$ is a parameter characterizing the diffusion of accelerated particles (G. Auriemma, T.K.G. & P. Lipari, in preparation), are contained in the magnetic field of the shell and interact with the shell material. Neutral secondaries freely leave the shell. These include neutrons with energy ≥ 1 TeV, which do not decay before leaving the supernova shell. If a young pulsar is born in SN1987A and protons are accelerated to ultra-high energies, the phenomenon might be observed for the first time in very high energy and ultra-high energy (UHE) γ -rays and upward-going, neutrino-induced muons.

Although the acceleration might be (and is indeed expected to be) quite different for the two classes of objects, we can compare the UHE cosmic ray luminosity of SN1987A to that of the well-studied X-ray binary system Cyg X-3, another candidate for UHE proton acceleration^{7,8}. The TeV γ -ray flux from Cyg X-3 is $4 \times 10^{-11} \text{ TeV/E particles cm}^{-2} \text{ s}^{-1}$, which corresponds to a source proton luminosity of $\sim 10^{39} \text{ erg s}^{-1}$. Such a luminosity could also be accommodated in supernovae containing pulsars with period smaller than 15 ms according to equation (1). If a young pulsar with $P \sim 1$ ms resides in SN1987A, luminosities of $(15/1)^4 = 5 \times 10^4$ Cygnus luminosity could be achieved for a duration of ~ 1 yr. The relative loss in lumi-

nosity from the fact that the Large Magellanic Cloud (LMC) is five times the 12 kpc Cyg X-3 distance is offset by the fact that the Cygnus accelerator has a duty cycle of only a few per cent⁸. Experiments are contemplated (one is already running in New Zealand) to observe this dramatic emission. For this experiment (JANZOS) to observe SN1987A at 10^{17} eV, it is sufficient to have a rate excess of about 10^3 over Cygnus luminosity, that is, $\sim 10^{42} \text{ erg s}^{-1}$. We point out here that this much power is unlikely to be realized because the supernova luminosity in the 10^{15} – 10^{17} eV energy range is constrained by the continuous cosmic ray (CR) flux at Earth.

An upper limit on the supernova luminosity in this range can be obtained from the assumption that the total CR flux between 10^{15} and 10^{17} eV is accelerated in young supernovae in our Galaxy. The observed energy flux of such particles at Earth is $7 \times 10^{-7} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$, which leads to a total energy content in the Galaxy of $\epsilon_{\text{CR}} = 3 \times 10^{-16} V_{\text{cont}} \text{ erg}$, where V_{cont} is the effective containment volume of CR in our Galaxy.

The total contribution of young supernovae to the pool of galactic CR between 10^{15} and 10^{17} eV is $\epsilon_{\text{YSN}} = \nu \tau_{\text{cont}} \Delta t f_n L_p$, where ν is the frequency of supernovae, τ_{cont} is the containment time of high-energy CR in the Galaxy, Δt is the time during which a typical supernova injects particles into the Galaxy, f_n is the fraction of the accelerated particles that are converted to neutrons with energies $> 10^{15}$ eV, and L_p is the luminosity in accelerated protons $> 10^{15}$ eV. The fraction f_n depends somewhat on the proton energy spectrum but is ~ 0.1 . We estimate $\tau_{\text{cont}} = 10^5$ yr (ref. 9), $\nu = 0.03 \text{ yr}^{-1}$, $\Delta t = 3$ yr, and $V_{\text{cont}} = 10^{67} \text{ cm}^3$. Then equating ϵ_{CR} and ϵ_{YSN} gives an upper limit $\nu_{30} L_p \leq 10^{41} V_{67} / \tau_5 \text{ erg s}^{-1}$. Here ν_{30} , V_{67} and τ_5 are the respective quantities in terms of our chosen estimated values.

There are large uncertainties in these estimated numerical values. Nevertheless the observed galactic CR spectrum places useful upper limits on the product of frequency and power of young supernovae that accelerate particles to energies $> 10^{15}$ eV. Clearly, only a small fraction ($< 1/100$) can do that with $L_p = 10^{42} \text{ erg s}^{-1}$.

Indeed if this acceleration mechanism works we can turn the argument around and relate the observed CR spectrum above 10^{15} eV to the distribution of initial pulsar periods and magnetic field strengths. This will be a subject of a future paper.

Some other observations can be used to place indirect limits on the power L_p that might be expected from SN1987A. From the fact that no anisotropy of cosmic rays (due to photons) is observed in the direction of the galactic centre one can put a

limit ~ 10 times Cyg X-3 luminosity during the 3 yr period of the experiment¹⁰. A somewhat weaker upper limit ($L_p \leq 10^{41} \text{ erg s}^{-1}$) is implied by the upper limit on neutrinos (upward muons) in deep underground experiments¹¹ from the direction of the centre of the Galaxy. The neutrino limits apply for a period of 1.5 yr during 1983–84. Recent limits on the pulsar contribution to the optical light curve of SN1987A (W. D. Arnett & S. Woosley; talks at George Mason University Supernova Workshop, October 1987) correspond very roughly to $10^{40} \text{ erg s}^{-1}$, assuming 10% efficiency for converting particle luminosity into visible light.

A luminosity of up to $10^{40} \text{ erg s}^{-1}$ in protons from SN1987A would be quite consistent with all these limits. If the output of SN1987A in ultra-high energy protons is of this order of magnitude, existing VHE and UHE γ -ray detectors in the Southern hemisphere will still be able to see fluxes from the supernova in the LMC provided the spectral index of the accelerated cosmic rays is < 2.4 . Were such an event to happen in the galactic centre, UHE γ -ray fluxes will of course be stronger by a distance factor of $(50/8)^2 \approx 30$.

We thank A. Harding for discussions. This research was supported by the University of Wisconsin Research Committee, with funds granted by the Wisconsin Alumni Research Foundation, by the US Department of Energy and by the National Science Foundation.

T. K. GAISSER & T. STANEV
Bartol Research Institute,
University of Delaware, Newark,
Delaware 19716, USA

F. HALZEN

Physics Department,
University of Wisconsin,
Madison, Wisconsin 53706, USA

1. Lagage, P. O. & Cesarsky, C. J. *Astr. Astrophys.* **125**, 249–257 (1983).
2. Sato, H. *Prog. Theor. Phys.* **58**, 549–559 (1977); *Mod. Phys. Lett.* **A2**, 801 (1987).
3. Berezhinsky, V. S. & Prilutsky, O. F. *Astr. Astrophys.* **66**, 325–331 (1978).
4. Berezhinsky, V. S. & Ginzburg, V. L. *Nature* **329**, 807–809 (1987).
5. Gaisser, T. K., Harding, A. K. & Stanev, T. *Nature* **329**, 314–316 (1987).
6. Gunn, J. E. & Ostriker, J. P. *Phys. Rev. Lett.* **22**, 728–731 (1969).
7. Hillas, A. M. *Nature* **312**, 50 (1984).
8. Protheroe, R. *Proc. Moscow Cosmic Ray Conference* (in the press).
9. Honda, B. *et al. Proc. 19th International Cosmic Ray Conference*, La Jolla, Vol. 2, p. 276 (1985).
10. Clay, R. W. *et al. Nature* **309**, 687–688 (1984).
11. Alkkofer, O. C. *et al. Astrophys. J.* **291**, 468–470 (1987).
12. Svoboda, R. *et al. Astrophys. J.* **315**, 420–424 (1987).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references. □

Thought of Dr Leggett

Sudip Chakravarty

The Problems of Physics. By A. J. Leggett. Oxford University Press:1987. Pp.192. Hbk £15, \$17.95; pbk £4.95, \$10.95.

To those who have followed Leggett's work professionally, this book will not come as a surprise. It is written in the style which is so characteristic of the author: complex conceptual questions are broken up, inessential pieces are removed and the ideas are finally presented with a simplicity that no one can imitate. Nonetheless, whilst reading the book I could not help asking myself how it would be received by those less familiar with Leggett's writings. I shall return to this point later. For the moment, it seems to me that no matter how one is swayed by one's own personal prejudices, the book is undeniably an exquisite presentation of frontier areas of contemporary physics. The questioning attitude can be a vital source of inspiration to young and would-be physicists. We all know that strong pressures to conform to existing paradigms come much too early in our lives.

Elementary particle physics and cosmology are currently undergoing revolutionary changes. Quantum chromodynamics has ushered in a new era in which the question "What are things made of?" has reached a new level of simplicity. We have learnt that the fundamental interactions owe their form to deep concepts such as gauge invariance. We have also learnt that two of those interactions, the electromagnetic and the weak, can be unified, which has led to spectacular predictions that have been experimentally well verified. To explain these developments in a language that is understandable to a non-expert needs an unusual amount of insight, and I suspect that the clarity with which they are discussed in this book will be welcome by all.

A point that Leggett makes over and over again, and one that is worth more than a pause, is that in physics we are continually making extrapolations. Often the laws that are tested firmly on a given scale are assumed to hold on scales which are unimaginably different. Leggett illustrates the point with an almost poetic poignancy: "If we imagine a microbe confined to the surface of a microscopic speck of dust floating in the middle of St Paul's Cathedral, the microbe's problem in inferring the properties of the Cathedral, or even the earth as a whole, would be trivial compared with ours".

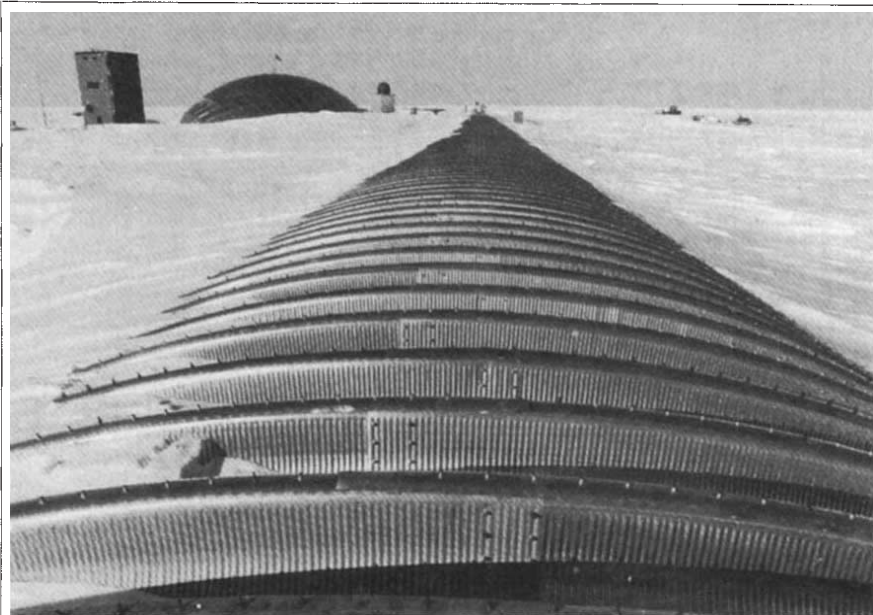
Here, of course, he means the reality that a cosmologist faces. But then did not Newton also make such an extrapolation

when he successfully described the motion of the planets? How are we able to make such extrapolations in physics? What guides us? What are the tools of present-day cosmology? How can we describe the first 10^{-35} seconds after the Big Bang? It almost sounds fantastic to a layman, and yet cosmologists are making steady progress in finding answers to these questions. New concrete concepts and tools of particle physics are being used successfully. The idea of broken symmetry, borrowed from condensed-matter physics, is exploited to propose how, as the Universe expanded, and cooled, the various symmetries were progressively broken, and the Universe evolved to the state that we live in. The scheme, however grand and elegant it may sound, is not without its difficulties. Almost all versions of this scenario predict unobserved debris of the Big Bang — monopoles, cosmic strings and so on. It is only recently that those pursuing the concept of an inflationary universe have begun to come to terms with such difficulties.

Leggett then moves on to physics on an entirely different scale, a scale that we are familiar with in our everyday life, the so-called condensed-matter physics. The subject matter in this area is extraordinarily diverse, so some important develop-

ments are sampled. In this respect I sometimes wonder how old-fashioned the textbooks are, and how naive the views of some of my colleagues are as to what condensed-matter physics is. Leggett has strived marvellously to dispel some of the antiquated notions in common currency. The sample that he provides is truly representative of the modern developments, and even experts will find the exposition insightful. But what is perhaps more important is the underlying view that the author holds. As he says, it is not uncommon to hear "that there are no new laws of nature to be discovered by studying condensed matter as such, since all behaviour of such matter follows, in principle, from the behaviour of its atomic or subatomic constituents; and second if this is so, then the study of complex matter cannot be as 'fundamental' as the study of the constituents themselves — indeed, that it is really rather a trivial occupation by comparison".

Leggett then goes on to expose the fallacious nature of this view, and in this in my opinion he has done a much-needed service to the community. His view that the most important advances in this area come about by the emergence of *qualitatively* new concepts is a point that even the practitioners of the field sometimes fail to recognize. Some of these concepts such as voltage, temperature, entropy and so on are so ingrained in our education that we fail to recognize their true role in the development of the field. Others, such as broken symmetry and universality classes, have found such wide applications that some of us tend to forget where they came from. I leave it as an amusing exercise to



Hidden depths — the Amundsen-Scott Station at the South Pole. The picture is taken from International Research in the Antarctic, which is based on reports from working groups of SCAR, the Scientific Committee on Antarctic Research. Published by the International Council of Scientific Committees with Oxford University Press, the book costs £25, \$39.95.

the reader to trace, for example, the origin of what is now known as the Higgs mechanism. The author also correctly notes that, when it is a matter of many interacting degrees of freedom, we often run into totally unexpected effects which are difficult to predict ahead of their experimental discovery, such as certainly been the case for superconductivity and superfluidity. And how about the newly discovered high-temperature superconductors? Those who believe that condensed-matter physics is an arid occupation of simply solving many-particle Schrödinger equations should take a lesson.

As some of us know, Leggett has been interested in the problem of quantum measurement for some time. Indeed, in recent years, he has played a major role in inspiring condensed-matter experimentalists to perform a number of beautiful experiments to test quantum mechanics on a macroscopic scale. It therefore does not come as a surprise to see a large amount of space devoted to the subject. In fact, the whole book betrays an uneasiness about quantum mechanics. However unorthodox the point of view of the author may be, I am sure that any serious reader will find his discussion of Bell's inequality delightfully simple and yet forceful. The same is true for his account of the 'cat paradox'. The question that he wishes us to consider is "whether, and if so where and how quantum mechanics breaks down in the face of increasing complexity..."

However, there are certain aspects of the book that I found disturbing. First, there is a tendency throughout the text to throw out a few questions, for example "why should nature seem to be described by Lagrangian field theory, and therefore submit herself to the stringent constraints imposed by it? Is the formalism universally valid, in fact, even when it gives rise to severe paradoxes?... Or are these questions themselves 'meaningless'?" Although these questions are not in themselves 'meaningless', they seem to lose their impact if they are not seen in context.

My second concern is to do with an unnecessary defensiveness, which perhaps stems from a particular type of sociological tension in the physics community today. Thus Stephen Hawking may proudly announce that the end is perhaps in sight for theoretical physics, but we know, all too well, what the fate of such prophecies is. We do not need to put up a straw man to conclude that "... far from the end of the road being in sight, we are still, after three hundred years, only at the beginning of a long journey along a path whose twists and turns promise to reveal vistas which at present are beyond our wildest imagination". □

Sudip Chakravarty is an Associate Professor in the Department of Physics, State University of New York at Stony Brook, Stony Brook, New York 11794, USA.

Climatological quirks

A. Slingo

A Climate Modelling Primer. By A. Henderson-Sellers and K. McGuffie. Wiley: 1987. Pp.217. £28.50, \$53.95.

CLIMATE modelling is very much in the news at present. Our understanding of the physical mechanisms responsible for such important natural variations as the ice ages and the El Niño/Southern Oscillation has improved through advances in both observational techniques and numerical modelling. At the same time, models have been used to predict the systematic changes expected from increasing concentrations of greenhouse gases liberated by anthropogenic activities, as well as the short-term effects of a nuclear war.

There is clearly a need for an introductory text which describes climate models and their application to a wide range of problems. The authors of this book have identified a gap in the literature, but their account is far too idiosyncratic to serve as a useful introduction and in many places it is actually misleading. For example, what little information there is on the physical basis of climate is scattered through the text, so the reader is in no position to understand the material on feedback processes and climate change which occupies much of the first chapter.

Three-dimensional models form the basis of serious climate research, yet they receive less attention than much simpler models. As justification, there is the extraordinary and incorrect assertion that in 1986 there was "a series of occurrences of apparently correct results being generated for the wrong reason".

Some quirks are simply juvenile, such as the attempts at humour in the dedication and acknowledgements, and the glosary definition of infinite as "quite a lot — really an awful lot". Others are more serious. There is a tendency to reproduce material from other sources without explanation, as in the section on cloud prediction in Chapter 6. In this and other places, the reader is swamped by unnecessary detail, while the underlying physics remains unexplained. Such detail does, however, lead to a useful bibliography.

The book is summed up by the 'spaghetti diagram', which purports to show the many interactions between elements of the climate system and which leaves the boxes blank for readers to fill in for themselves. As an alternative text, I would recommend *An Introduction to Three-dimensional Climate Modeling* by Washington and Parkinson, which, although its remit is more restrictive, does at least have the stamp of authority which comes from authors who are acknowledged experts in this field. □

A. Slingo, Meteorological Office, Bracknell, UK, will shortly be moving to the National Center for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307, USA.

Static action

Peter Dunnill

Immobilized Cells: Principles and Applications. By J. Tampion and M.D. Tampion. Cambridge University Press: 1987. Pp.257. £30, \$59.50.

Immobilised Enzymes and Cells. By A. Rosevear, J.F. Kennedy and J.M.S. Cabral. Adam Hilger, Bristol/Taylor & Francis, Philadelphia: 1987. Pp.248. £37.50, \$102.

IT IS now 80 years since Michaelis of enzyme kinetics fame described the adsorption of an enzyme to charcoal, and 30 years since the systematic study of enzyme immobilization for practical application began in earnest. Success with attachment of enzymes to supports encouraged studies of whole-cell immobilization, though immobilized cells had been used for a century, largely unconsciously, in fields such as sewage treatment and vinegar production.

There are now dozens of methods of enzyme and cell immobilization and any guidance on appropriate procedures is

very welcome. As their titles imply, these two books take quite different approaches, with one focusing just on immobilized cells and the other considering both enzymes and cells. For newcomers to applied biocatalysis, the latter approach is particularly valuable because the first decision to be made is whether to use whole cells or enzymes.

That choice might be more straightforward if scientists could agree on what they mean by the use of the word 'cell'. Thus on page 1 Tampion and Tampion define an immobilized cell as "a cell or remnant thereof that by natural or artificial means is prevented from moving independently". Aside from cell organelles, however, a remnant of a cell effectively functions as an enzyme catalyst lacking the organization needed for, say, co-factor recycling. Rosevear *et al.* refer to such systems as "dead cells", but the European Federation of Biotechnology Working Party disapproves of this term. Equally it dislikes terms such as "living", but welcomes "viable", "non-growing" and "respiring" in appropriate circumstances. At present we don't understand why immobilized cells are often much more stable as catalysts than free cells, so

precision in publications describing their state is vital.

Both books review the published methods of immobilization but the volume by Rosevear, Kennedy and Cabral is explicitly a "laboratory book", giving guidance to the student or research worker who needs to choose an appropriate method. The Tampions' volume mixes principles, methods and applications, its target audience is less clear and the exclusion of enzymes is justified in rather odd terms, thus: "The interest in immobilized cells draws attention to the fact that due to economic or scientific reasons there are drawbacks to the use of comparable immobilized enzyme systems". However, the "classical" example of industrial application cited — high-fructose corn syrup production — uses principally immobilized enzyme and not immobilized cell technology, as the book indicates. Rosevear and his colleagues recognize this but state the tonnage a million-fold low at three tonnes per year in North America.

Industrial research on immobilized enzymes has fallen back from a peak because it failed to address fully the key issue, which is that many enzymes are still not sufficiently stable for practical application even when immobilized. It seems likely that cell immobilization studies as currently carried out will also falter because often the cells do not perform well for long enough. Neither book sets out to address these issues or the possible solutions such as protein engineering for enzymes and genetic engineering for cells. Protein engineering techniques already show promise in the preparation of more stable enzyme catalysts, and in principle we have the knowledge to alter the genetic material of a cell so as to separate synthetic capabilities from reproductive ones. The researcher must bear in mind that these newer techniques, and the recent development of better synthetic membranes for retaining biocatalysts rather than immobilizing them, may be of great importance in the future.

To be fair, the present authors would have found little literature to review on these innovations. Both books make a valuable contribution to the business of keeping up with one of the main techniques of biotechnology, and, if the intrinsic properties of free enzymes and cells can be improved, they will help to usher in a new era of applied biocatalysis. □

Peter Dunnill is a Professor in the Department of Chemical and Biochemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK.

• Just published by Elsevier Applied Science is *Bioreactor Immobilized Enzymes and Cells: Fundamentals and Applications*, edited by Murray Moo-Young, a collection of papers covering the full range of cell types along with applications of the systems. Price is £39.

Capital letters

Colin Patterson

Simple Curiosity: Letters from George Gaylord Simpson to His Family, 1921–1970. Edited by Léo F. Laporte. *University of California Press: 1987. Pp. 340. \$29.95, £16.70.*

GEORGE Gaylord Simpson (1902–1984) was identified as "the greatest paleontologist since Cuvier" by the supervisor of his PhD, and is referred to as "the Darwin of the twentieth century" in the blurb of this book. Neither opinion stretches belief. Simpson was a prodigious worker: in 60 years, he published hundreds of technical



G. G. Simpson — an exceptionally gifted youth.

papers and over 20 books, among them two personal memoirs. The first was *Attending Marvels: A Patagonian Journal* (1934), the second *Concession to the Improbable: An Unconventional Autobiography* (1978).

Attending Marvels is a wonderful travel book, the account of GG's first fossil-collecting expedition to Patagonia in 1931–1932. It is full of the fire and joy of an exceptionally gifted youth. *Concession to the Improbable*, GG's autobiography, though still characterized by style and wit, left a sour taste with many readers; inevitably, it is an old man's book. *Simple Curiosity* fills some of the gap between those two memoirs. The letters collected here were written to his sister Martha (1898–1984) and to his parents. Martha had deposited about 150 family letters in the American Philosophical Society, and others were found among Simpson's papers after his death.

In Simpson's autobiography, his youth,

up to his return from Patagonia in 1931, occupies hardly 10 per cent of the text. Happily, more than half of this new book covers the decade 1921–1931, and great fun it all is. During that time, he studied at the University of Colorado and at Yale; married, had four daughters, and separated; spent a year in London, a summer in Paris, a year in Argentina, and travelled widely in Europe and North America; he learned French, German, Spanish and Egyptian (hieroglyphs); published about 70 scientific papers and two books; became associate curator of fossil mammals at "the greatest Natural History Museum in the world" (the American Museum of Natural History); and detailed his delight in all this, principally in the letters to his sister.

The title of the book, *Simple Curiosity*, must have been suggested to Laporte by its recurrence in two letters to Martha (1926 and 1927) on the motives of the scientist. These early letters are full of lively commentary on ideas and events, with much high-class whimsy, some graphic, some polyglot, some in doggerel. One 1929 letter begins with three paragraphs that would sit well in John Lennon's *In His Own Write* before switching into French (translated by the editor). Léo Laporte has done a graceful job as editor, splitting the letters chronologically into a dozen more-or-less coherent sections, introducing each by a commentary and an apt photograph, and providing footnotes just where one needs them.

My only dissent from Laporte's editorship (admittedly a major one) concerns the last third of the book. On returning to New York from Patagonia in 1931, Simpson filed for separation from his first wife, and in 1936, after another absence of almost a year in Patagonia and Europe, sued for divorce. The long and unpleasant suit chronicled in the letters finally ended in 1938, and within a month Simpson married the psychologist Anne Roe, who had grown up with him in Denver. They remained together for 46 years (and in the 1960s became the first married couple each to hold full professorships at Harvard).

For whatever reason, after the marriage in 1938 the spark goes out of Simpson's letters to his family and they hardly record more than comings, goings and thank-yous. The wartime letters (Simpson was in North Africa and Italy for 18 months in military intelligence) are constrained by censorship and format, and after the war, apart from a couple of pages of hilarity on the organization of Harvard, there is little beyond the family album. If Laporte had stopped short in 1938, we should have had a smaller book, but a more remarkable portrait of a most remarkable man. □

Colin Patterson is in the Department of Palaeontology, British Museum (Natural History), London SW7 5BD, UK.

Explanations for aberrations

K.W. Jones

Cancer Cytogenetics. By Sverre Heim and Felix Mitelman. *Alan R. Liss, New York/Wiley, Chichester: 1987. Pp. 309. \$35, £27.50.*

FOR many years, the study of chromosomes of tumours seemed to offer little insight into the causation of cancer and attracted comparatively little general interest. Things changed with the original finding of the Ph1 marker chromosome in chronic myeloid leukaemia by Nowell and Hungerford in 1961. Since then, improvements in cytological techniques and the discovery of viral oncogenes and cellular proto-oncogenes has made cancer cytogenetics a focal discipline for molecular biologists, virologists, geneticists, pathologists and clinicians.

Given the multiplicity of contributing subjects, and the growing appreciation of the complexities underlying the pathogenesis of cancer, this is a difficult field to encompass in a single book. Heim and Mitelman have nonetheless succeeded in producing a readable, critical and up-to-date account, which not only conveys both historical perspective and the current excitement in this area, but which will be a valuable reference source.

The book emphasizes both the successes and limitations of cancer cytogenetics, for example the fact that most tumours have remained refractory to cytogenetic analysis so that 87 per cent of analysable cancers are leukaemias. Unfortunately, the intractable class includes carcinomas which are quantitatively the most important. Moreover, as the authors point out, genetic changes of as much as 10^6 – 10^7 base pairs may occur without there being a visible cytogenetic consequence, so that chromosome studies, although invaluable, reveal just the tip of what might be a very large iceberg.

Nevertheless, the striking fact which has emerged is that chromosomal aberrations in human cancer are not random but are clustered in approximately 83 chromosomal bands; at least one of these particular regions, apparently, is involved in 96 per cent of neoplasms. Some of these are known fragile sites and/or correspond to the locations of known proto-oncogenes, which perhaps explains why changes in these abnormal chromosome patterns frequently have clinical significance both as to treatment and prognosis. Interestingly, the precise character of abnormalities also appears to show geographical variation in apparently identical malignancies, suggesting differences in relevant gene frequencies or in exposure to

oncogenic factors.

Cancer Cytogenetics is as much a clinical manual as an academic text, with specific sections in each chapter summarizing the clinical correlations for various cancers, as one would expect from two distinguished clinical geneticists. But because this material is skilfully integrated with accounts of the underlying molecular mechanisms, where such information exists, the book will draw molecular biologists and clinicians into a closer mutual understanding, which, I suspect, was the authors' intention. The multifaceted approach to the subject matter constantly throws up stimulating questions and means the reader's interest is held throughout what, in other hands, could easily have been a dull book. The authors are not reticent in putting forward

their own views, but fact, uncertainty, hypothesis and opinion are carefully and consistently differentiated, which makes the volume of particular value to students.

In all, this is a stimulating and lucid book. It contains some 14 tables, which condense and classify a large amount of useful information, 63 figures and diagrams, and over 600 references, many of which are to papers published in 1987. It will appeal to oncologists and molecular biologists actively engaged in research, as well as to pathologists, clinicians, teachers and students. Given its wide appeal, in my view no science library can afford not to have several copies of it. □

K.W. Jones is a Reader in the Department of Genetics, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JN, UK.

Versatile vectors

Tim Harris

Vectors: A Survey of Molecular Cloning Vectors and Their Uses. Edited by Raymond L. Rodriguez and David T. Denhardt. *Butterworths: 1988. Pp. 578. £45, \$54.95.*

ONCE it was that one obtained a vector for cloning DNA from a friend or colleague. These days, as molecular biology kits (and molecular biologists) proliferate, many vectors can be bought off the shelf, tailor-made for a specific task, be it cDNA or genomic cloning, transcription or expression, or for introduction into animal or plant cells. The diversity and usefulness of modern vectors can be appreciated from this collection of multi-author reviews, although, inevitably, some recent advances have not been included.

Balance is all-important in a book of this kind; to some extent judgement over which topics should have been covered in more detail and which in less detail is a matter of individual preference. For me, the editors have not got it quite right. The emphasis is on bacteria, with at least two-thirds of the chapters dealing with both the gram negative and gram positive types. General topics include a useful potted history and genealogy of pBR322, and descriptions of M13 and lambda phages, cosmids and phagemids all at a level suitable for third-year undergraduates. More specific coverage of plasmids for cloning promoters and terminators, SP6 and T7 promoter-directed transcription, copy number control and for cloning into broad host ranges is also provided in this section. There is the notable absence, however, of a chapter dedicated to site-directed mutagenesis — an increasingly important technique which is only referred to in passing, in amongst the

filamentous phages.

The remainder of the book is concerned with vectors for other systems, including fungal cells, insects and insect cells, and animal and plant cells. It is not quite fair to say that these chapters seem to have been added as an afterthought, but there is a suspicion of it — quite a surprise considering the importance of the subject matter for molecular biology in general and biotechnology in particular. Mammalian cell vectors are particularly poorly dealt with, little attempt being made to evaluate critically the relative merits of one system versus another, and there is little mention of the vectors being used to produce large amounts of proteins for therapy (for example t-PA, erythropoietin and various antibodies).

It is a pity that an industrial viewpoint is not given here, as it is in one of the two chapters concerned with the transformation of plant cells. A similar criticism can be levelled at the section on yeast, which although it covers the different types of shuttle plasmid available, does not go into the details of those used to express foreign genes such as hepatitis B surface antigen or α 1-antitrypsin. While application is not the be all and end all, a lot of people do use vectors for expression purposes.

So, like the curate's egg, the book is good in parts. If bacterial vectors are what you want to know about, then it will be pretty useful, although the quality of the diagrams in general leaves much to be desired. For other cells it would be better to go to the excellent, up-to-date and relevant reviews that are available. □

Tim Harris is Head of Molecular Biology at Celltech Ltd, 216 Bath Road, Slough SL1 4EN, UK.

New in Britain

Invisible Frontiers: The Race to Synthesize a Human Gene, by Stephen S. Hall, is now available in Britain. Publisher is Sidgwick and Jackson, price is £13.95. For review see *Nature* 329, 399 (1987).

Evidence of bias in estimations of earthquake size

Göran Ekström* & Adam M. Dziewonski

Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

Magnitudes of earthquakes determined using traditional methods show systematic deviations, dependent on tectonic setting, from accurate estimates of earthquake size. Magnitudes are overestimated for most of the continental earthquakes, underestimated for mid-oceanic ridges and differences of varying sign are obtained for subduction zones. This is important in evaluating the seismic risk or consideration of effects such as the seismic versus aseismic strain release as seismic moments determined from magnitudes can be wrong by as much as a factor of four.

UNTIL 1935 when C. F. Richter¹ first proposed a general magnitude scale the size of an earthquake was primarily described by the damage it caused. This, of course, depends on a great many subjective factors. The advantage of Richter's scale is that it is a simple function of the amplitude of ground motion measured on seismographic recordings at any distance from the earthquake source. During the past 50 years several variations of the magnitude scale were developed. For the large, shallow and most destructive earthquakes the surface wave magnitude M_s is most important. One commonly used definition² is

$$M_s = \log A_{20} + 1.66 \log \Delta + 2.0$$

where A_{20} is the amplitude in micrometres of surface waves with a period of ~ 20 s, and Δ is the distance in degrees between the earthquake and the station.

A severe limitation of the magnitude scale is its purely empirical nature: there is no direct relationship to the physics of the earthquake process or the effects on the Earth's structure, or the excitation of seismic waves. The progress in seismological theory, digital instrumentation and the availability of high-speed computers now makes it practical to determine routinely the seismic moment, a different measure of earthquake size. One definition of the seismic moment is³:

$$M_0 = \int_{\Omega} \mu u \, d\Omega$$

where μ is the shear modulus, u is the relative displacement of one side of the fault with respect to the other, and Ω is the fault area.

Here we present the results of an investigation on both global and regional scale of the empirical relationship between the surface wave magnitude M_s as reported by the service agencies, and the scalar seismic moment M_0 . The importance of investigating this relationship lies primarily in that direct determination of the seismic moment can be performed only for large numbers of recent earthquakes. Very few seismic moment determinations have been made for events which occurred before 1977, and a magnitude estimate is often the only parameter available in earthquake catalogues. These magnitude determinations may have important practical applications; for example, in estimates of seismic risk and setting of local building codes. It is important, therefore, to compare these estimates with a much more objective measure of earthquake size.

Quantification of earthquakes

The first station of the two pioneering digital global seismographic networks was set up in the mid-1970s. Both the International Deployment of Accelerometers (IDA)⁴ and Global Digital Seismographic Network (GDSN)⁵ achieved full strength

by 1980. The IDA network, consisting of LaCoste-Romberg gravimeters, was designed to study phenomena at ultra-long periods (several hundred to several thousand seconds) and is primarily useful in analysing the waves generated by large ($M_s \geq 6.5$) earthquakes⁶. The GDSN network was, on the other hand, designed to have peak performance in the frequency band used in discrimination between earthquakes and nuclear explosions (0.04–1 Hz). Most earthquakes with magnitudes > 5.2 – 5.4 are recorded by the GDSN with a satisfactory signal-to-noise ratio.

The centroid-moment tensor (CMT) technique was developed by Dziewonski *et al.*⁷ to take advantage of these data to study, on a systematic basis, moderate-size earthquakes. The inverse problem is formulated such that the synthetic seismogram, calculated to determine the source parameters, should match, in the least-squares sense, the body-wave portion of a seismogram: from the P wave until, but excluding, the arrival of surface waves. This choice of data is intended to maximize the sensitivity to source radiation patterns and minimize the effect of lateral heterogeneities. The method was later generalized to incorporate very long period data (mantle waves)⁸. Also, procedures were developed for correction of synthetic seismograms and differential kernels for known lateral heterogeneity^{9,10}.

The seismic moment is a derived parameter of a complete CMT solution which generally consists of six independent elements of the moment tensor and the four spatio-temporal parameters of the best location of a point source: the source centroid. The seismic moment, M_0 , is determined by averaging the absolute values of the largest positive and largest negative eigenvalues of the moment tensor. For earthquakes with moments larger than 5×10^{24} dyn cm the errors due to random noise are negligible. The most important effect for shallow earthquakes ($h < 30$ km) is the trade-off between the seismic moment and the dip of a fault plane (δ). For a typical thrust earthquake, for example, the amplitudes of the radiated long-period waves depend primarily on the product $M_0 \sin 2\delta$, which, therefore, is well constrained in the inversion. An error in the estimated dip angle will, however, translate into a corresponding error in the scalar moment estimate. It can be expected that this trade-off will lead to random, rather than systematic, errors in M_0 . The comparison of seismic moments derived from separate inversions of the body-wave and mantle-wave data suggests that the relative standard error in determination of M_0 is 10–20%. This is about one order of magnitude better than the estimates of seismic moments based on M_s values.

The centroid-moment tensor (CMT) technique has been applied to nearly 6,000 earthquakes that occurred between 1977 and 1987^{11,12}. This data set provides uniform global coverage for earthquakes with a moment $\geq 2 \times 10^{24}$ dyn cm, although, in some cases, earthquakes with moments as low as 3×10^{23} dyn cm can be reliably analysed.

Of the commonly reported magnitudes, M_s is generally considered to correlate best with the seismic moment. The moment

* Present address: Lamont-Doherty Geological Observatory, Palisades, New York 10964, USA.

magnitude, M_w , introduced by Kanamori¹³ ($M_w = \frac{2}{3} \log M_0 - 10.73$) indeed agrees well with M_s for events larger than a magnitude of ~ 6.5 . The inverse of this relationship, therefore, provides a means of estimating M_0 when only M_s is known.

The uncertainty associated with such estimates is illustrated by consideration of the largest earthquakes to occur in 1985 and 1986. Table 1 lists all shallow-focus earthquakes that occurred in these years with either M_s or $M_w > 7.4$. Large discrepancies between M_s and M_w are apparent for several of the events, and an ordering according to surface-wave magnitude would not coincide with an ordering in terms of scalar moment. Most egregious is the situation for the 20 October 1986 Kermadec Islands' earthquake. A surface-wave magnitude of 8.2 is the largest magnitude reported by the NEIC (National Earthquake Information Center of the US Geological Survey) in the past ten years, and the 0.5-unit discrepancy with M_w corresponds to a difference of a factor of five between the observed seismic moment and the seismic moment that would be inferred from M_s .

If discrepancies such as this one are common in the catalogues of historical earthquakes, estimates of rates of seismic deformation, identification of seismic gaps and seismic-risk assessments are subject to large errors. It would be particularly serious if such discrepancies were systematic rather than random. The difference may be difficult to tell when the consideration is limited to small numbers of earthquakes. However, when larger numbers of events are studied, systematic differences between geographical regions become apparent. Figure 1 shows the global pattern of differences between the reported surface-wave magnitude and the derived moment magnitude M_w for 471 earthquakes with $M_0 \geq 2 \times 10^{25}$ dyn cm ($M_w \geq 6.2$). Two observations can be made. First, it is more common for M_s to be smaller than M_w than vice versa. Second, although the scatter is large, there is indeed some regional consistency with clusters of positive and negative residuals. The fact that more than two thirds of the residuals are negative suggests that M_w does not provide an adequate average relationship for this range of scalar moments. A different global average should, therefore, be determined before residuals are calculated and compared.

A global average relationship

Figure 2a shows the data set that was considered in the analysis. It consists of 2,341 reported M_s measurements from the Preliminary Determination of Epicenters (PDE) published by the NEIC, and corresponding scalar moments from the catalogue of source mechanisms determined at Harvard using the CMT method of analysis. Only earthquakes for which both the NEIC and the CMT depths are < 50 km are included. The moment magnitude M_w provides a good fit to the observations for earthquakes with $M_0 \geq 10^{26}$ dyn cm. For smaller moments M_w is larger than M_s and the trend of the observations follows a slope which approaches unity.

This is in agreement with theoretical considerations of the M_s measurement. The standard surface-wave magnitude is calculated from observations of the logarithmic amplitude of surface

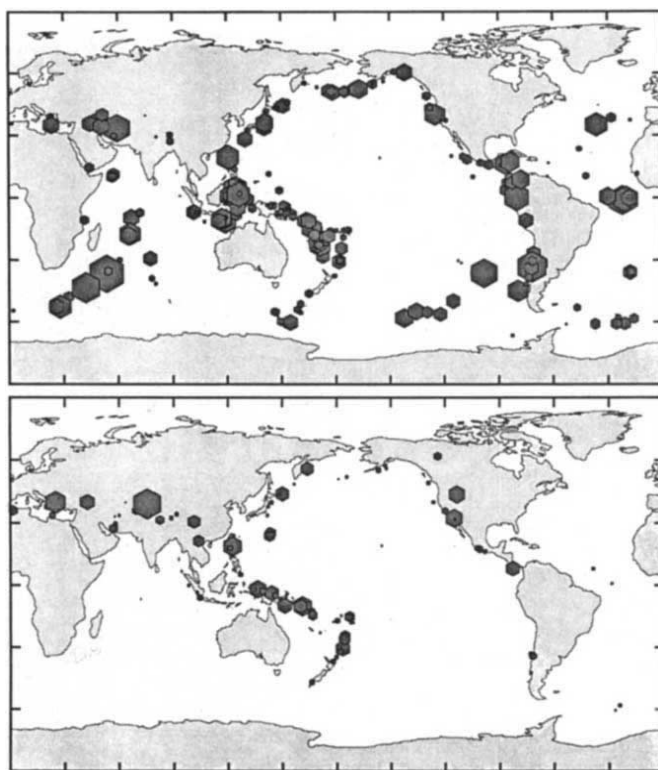


Fig. 1 Global maps showing the difference $M_s - M_w$ for shallow events ($h < 50$ km) with $M_0 \geq 2 \times 10^{25}$ dyn cm. Top panel shows events with negative deviations ($M_s < M_w$), bottom panel shows positive deviations. The size of each symbol is proportional to the absolute value of the difference and the largest symbol corresponds to a difference of 0.8 magnitude units.

waves with ~ 20 s period at one to several stations. For details of the measurement and averaging procedures, see Richter¹⁴ and Lienkamper¹⁵. If the moment release in the earthquake occurred instantaneously and at a single point in space, M_s would be proportional to $\log M_0$ with a proportionality factor of one, because the amplitudes of seismic waves at any period would be linearly proportional to M_0 . Due to the finite duration and extent of the earthquake source, the excitation of seismic waves falls off with increasing frequency, making the proportionality factor between M_s and $\log M_0 < 1$. That this factor is actually close to $2/3$ for a relatively wide range of scalar moments suggests that earthquakes, on average, follow source scaling relationships similar to those described by, for example, Kanamori and Anderson¹⁶.

Because of the observed and expected decrease in slope with increasing scalar moment, a linear relation will only provide an acceptable fit to a limited range of scalar moments. A least-squares-fitting straight line will in addition be dominated by more frequently occurring events of smaller moment (see refs 17 and 18). To remove the dominance of the smaller events, the observations were reduced by averaging observed M_s values in ranges of 0.1 units of $\log M_0$; the reduced data are shown in Fig. 2b. For the smallest events ($M_0 \leq 2 \times 10^{24}$ dyn cm) there is some scatter and the data set is biased towards large values of M_s both by the incomplete reporting of low-magnitude earthquakes and the selection of events to analyse with the CMT method. In the range 2×10^{24} to 2×10^{27} dyn cm the data points define a smooth curve with decreasing slope. For the largest events ($M_0 \geq 2 \times 10^{27}$ dyn cm) there is again some scatter due in part to the small number of observations, but also to the saturation of the surface wave magnitude for very large earthquakes.

To associate M_s with measurements of M_0 we introduce an empirical moment magnitude relationship, $M_s(M_0)$, which, for

Table 1 Comparison of surface wave magnitude as reported by the USGS and moment magnitude calculated from the seismic moment for eight shallow earthquakes in 1985 and 1986

Date	Location	M_s	M_w
3 March 1985	Central Chile	7.8	8.0
23 August 1985	Southern Xinjiang	7.5	7.0
19 Sept. 1985	Michoacan, Mexico	8.1	8.0
21 Sept. 1985	Michoacan, Mexico	7.5	7.5
7 May 1986	Andreanof Islands	7.7	8.0
14 August 1986	Molucca Passage	7.2	7.5
20 Oct. 1986	Kermadec Islands	8.2	7.7
14 Nov. 1986	Taiwan	7.8	7.4

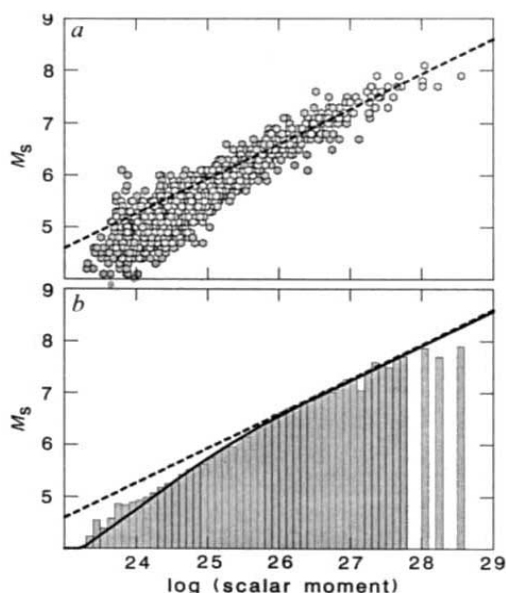


Fig. 2 The data considered in this study. *a*, Individual data points. *b*, Observations averaged in 0.1-unit-wide ranges of $\log M_0$. The dashed line corresponds to the moment-magnitude relation M_w of Kanamori¹³ and the solid line represents $\bar{M}_s$.

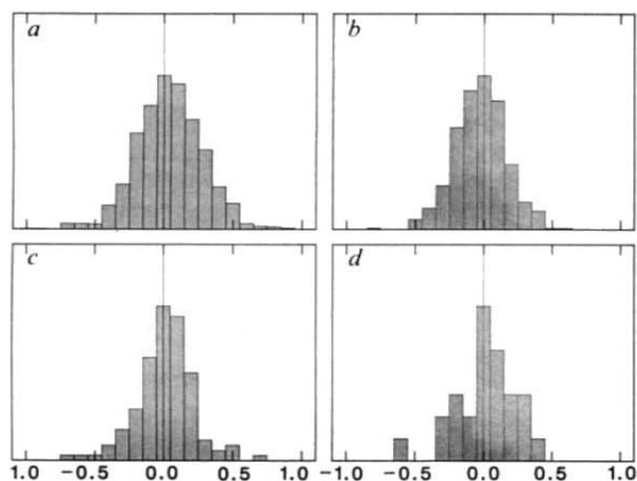


Fig. 3 Distribution of $M_s - \bar{M}_s$ residuals for four order of magnitude ranges of scalar moment: *a*, $10^{24} \leq M_0 < 10^{25}$ dyn cm, 1,376 events, $\sigma = 0.24$; *b*, $10^{25} \leq M_0 < 10^{26}$ dyn cm, 567 events, $\sigma = 0.19$; *c*, $10^{26} \leq M_0 < 10^{27}$ dyn cm, 127 events, $\sigma = 0.21$; *d*, $10^{27} \leq M_0 < 10^{28}$ dyn cm, 27 events, $\sigma = 0.22$.

each value of M_0 , provides a magnitude $\bar{M}_s$ which approximates the global average value of M_s observed for that scalar moment. Given the observed and theoretically predicted change in the relationship between M_s and $\log M_0$, from a slope of unity for small earthquakes to 2/3 for moderate to large events, a search is made for a simple relationship $\bar{M}_s(M_0)$ which is determined by three parameters a , b and k and has the form

$$\bar{M}_s = \begin{cases} k - \frac{(a+b)}{6} + \log M_0 & \log M_0 < a \\ k - \frac{(a+b)}{6} + \log M_0 - \frac{(\log M_0 - a)^2}{6(b-a)} & a \leq \log M_0 \leq b \\ k + \frac{2}{3} \log M_0 & \log M_0 > b \end{cases} \quad (1)$$

This corresponds to a linear relationship with slope one for $\log M_0 < a$, a second linear segment with slope 2/3 for $\log M_0 > b$, and a gradual transition in slope between a and b .

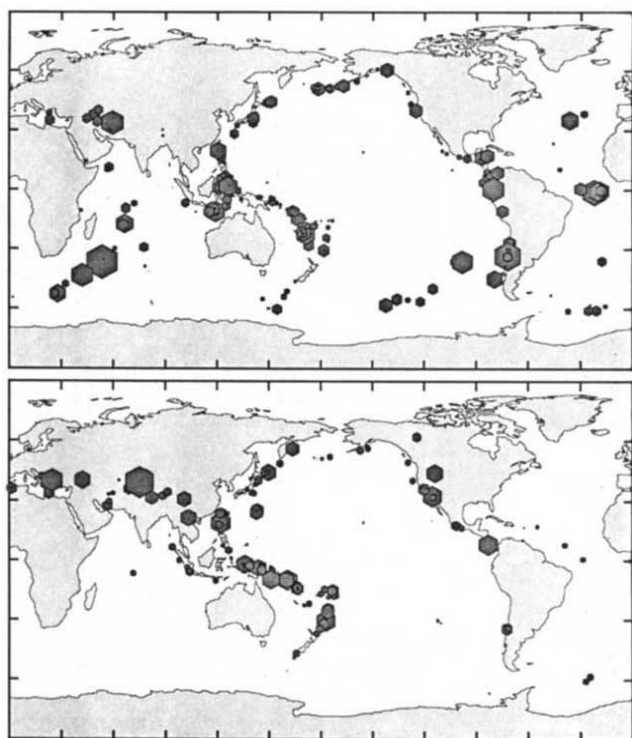


Fig. 4 Global maps showing the distribution of negative (top panel) and positive (bottom panel) residuals $M_s - \bar{M}_s$ for the same set of earthquakes as in Fig. 1. The size of each symbol is proportional to the absolute value of the residual and the largest symbol corresponds to a difference of 0.7 magnitude units.

A good fit to the reduced data for earthquakes with moments in the range 2×10^{24} to 1×10^{28} dyn cm is obtained with $a = 24.5$, $b = 26.4$ and $k = -10.76$ (which is nearly identical to $k = -10.73$ of ref. 13), and these values will define the empirical moment magnitude $\bar{M}_s$ in what follows. Using these values we can rewrite equation (1) in the form

$$\log M_0 = \begin{cases} 19.24 + M_s & M_s < 5.3 \\ 30.20 - \sqrt{92.45 - 11.40 M_s} & 5.3 \leq M_s \leq 6.8 \\ 16.14 + \frac{3}{2} M_s & M_s > 6.8 \end{cases} \quad (2)$$

which can be used to estimate the scalar moment from a measurement of M_s , subject to the uncertainties and biases explored here. Figure 2*b* shows the fit of $\bar{M}_s$ and M_w to the observations. The difference between M_w and $\bar{M}_s$ in predicting the scalar moment for an earthquake with $M_s = 5$ is a factor of 5, which is certainly significant. Figure 3 shows the histograms of residuals $M_s - \bar{M}_s$ for four different ranges of scalar moment. The fact that the distributions are close to symmetric and centred on zero shows that $\bar{M}_s$ can be used to predict M_0 by identifying it with $\bar{M}_s$ over a range of magnitudes from $M_s \approx 5$ to $M_s \approx 8$. Also, while there is no question that seismic moments determined from waveform analysis is a superior measure of earthquake size, the very large figures (of the order of 10^{26}) are cumbersome to use. Equation (1) provides the scaling to the much more convenient and familiar range of the magnitude scale.

Regional variations

Figure 4 shows the global distribution of negative and positive $M_s - \bar{M}_s$ residuals for events with $M_0 \geq 2 \times 10^{25}$ dyn cm. Several features stand out better than in Fig. 1. Events in Central Asia have consistently high reported M_s values as do events in the Aegean and continental events in North America. Low M_s regions include primarily ridges and fracture zones. Subduction

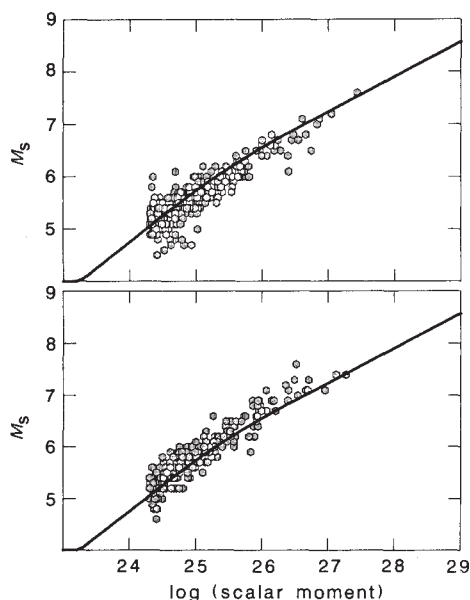


Fig. 5 Distribution of observations for *a*, ridge and fracture zones and *b*, continental earthquakes with $M_0 \geq 2 \times 10^{24}$ dyn cm. The solid line represents $\bar{M}_s$.

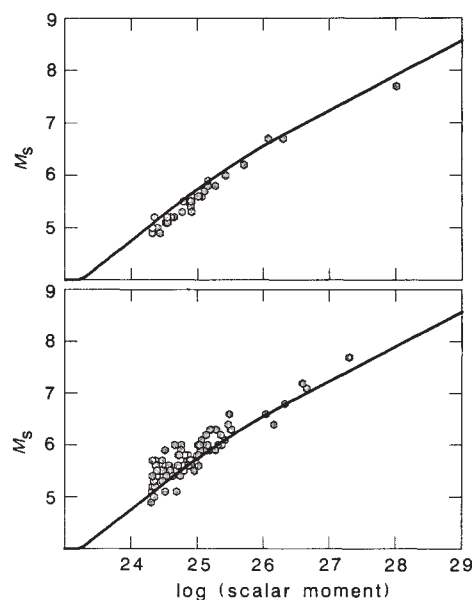


Fig. 6 Distribution of observations for thrust-type earthquakes in *a*, the Andreanof Islands region of the Aleutian Islands and *b*, the Tonga-Kermadec subduction zone.

zones do not show a simple pattern; the Aleutian arc, New Hebrides, and Chile are predominantly low while $\bar{M}_s$ is overestimated for Tonga-Kermadec.

It is now clear that the $(M_s - M_w)$ differences reported in Table 1 are not random, but are expressions of a systematic pattern. Even though we do not intend here to provide an explanation for the observed regional variations, it may be useful to examine the effect of possible factors.

The generation of 20 s surface waves depends on focal depth, faulting geometry, as well as the Earth's structure near the earthquake source¹⁹. If variations in these factors were a primary cause for the observed deviations, we would expect similar tectonic regions to exhibit a similar bias. To some extent this is also observed. Figure 5*a* shows the data points for globally distributed ridge and fracture zone events with $M_0 \geq 2 \times 10^{24}$ dyn cm and Fig. 5*b* shows data points for continental earthquakes. The difference in average reported $\bar{M}_s$ between these two tectonic regions for a moderate-size earthquake is ~ 0.4 magnitude units; this is equivalent to a factor of 4 in relative rate of moment release, for example. Other tectonic settings are not equally consistent. Figure 6 shows the distribution of data points for shallow-thrust earthquakes in two similar tectonic regions; the Aleutian Islands (between 180° and 170° W) and Tonga-Kermadec (between 40° S and 20° S) subduction zones. In spite of the similarities between these two settings, the deviations from $\bar{M}_s$ are of opposite sign.

A second factor which may contribute to the deviations is related to the propagation of the surface waves from the source region to the recording station. It was observed at an early stage²⁰ that the attenuation of 20-s surface waves was strongly path dependent, and attempts were made²¹ to compensate for this effect when calculating $\bar{M}_s$, both by consideration of the individual paths and by the introduction of station corrections. This practice is not currently used by the reporting agencies. Even though this factor may be an important one, it would fail to explain the opposite patterns of deviations observed for events in the Tonga-Kermadec and New Hebrides subduction zones, that are $< 2,000$ km apart and, therefore, are recorded by similarly distributed stations.

A third contributing factor is the combined effect of azimuthally varying radiation coefficients and non-uniform

station coverage²². Taken by itself it fails to explain the consistent, but opposite, patterns observed in continental and ridge-fracture zone areas, where both dip-slip and strike-slip events of different orientation are frequent.

In Fig. 7 $\bar{M}_s$ magnitudes for earthquakes in these two tectonic regions have been averaged in narrow ranges of $\log M_0$. In addition to the observation that $\bar{M}_s$ is greater for continental earthquakes over a wide range of scalar moments, the data also suggest that the difference between the two regions is not uniform but that it increases with scalar moment, except for the few very largest events. This may indicate that an additional factor contributing to the regional pattern are differences in source scaling relationships between different tectonic environments. Figure 7 shows that for continental earthquakes the change in slope from unity to a smaller value occurs at a larger scalar moment. This suggests that for a particular scalar moment, these earthquakes have a higher corner frequency and, therefore, also higher stress drop. This is consistent with what has been found by other investigators²³⁻²⁵.

We conclude that no single factor accounts for the observed regional pattern of $\bar{M}_s$ bias. Certainly it is possible that the $\bar{M}_s$ measurement could be improved by compensating for some of the known factors, such as the variable attenuation along

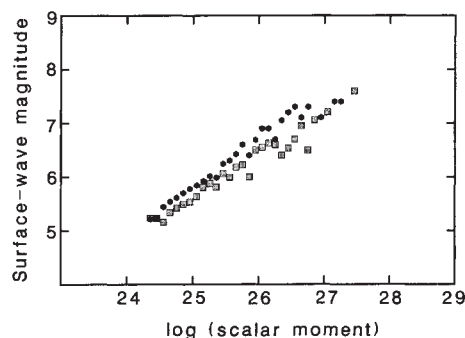


Fig. 7 $\bar{M}_s$ observations averaged in 0.1-wide intervals of $\log M_0$ for continental (filled hexagons) and ridge and fracture zone earthquakes (squares).

different paths, and by ensuring that an azimuthally uniform coverage of stations is used in the averaging calculation. To compensate for other factors, such as focal depth, fault geometry and corner frequency would require such a detailed knowledge of the earthquake source that the M_s measurement itself would be redundant.

The results of this analysis can be summarized in five points.

(1) A global average moment-magnitude relationship $\bar{M}_s$ has been defined which can be used to predict M_0 over a wide range of magnitudes and scalar moments.

(2) The variance of surface wave measurements for an event of a particular scalar moment is ~ 0.2 magnitude units.

(3) Large regional biases in M_s exist.

(4) Differences in source scaling may explain some of the differences. Specifically, observations show that the transition from a slope of unity to a smaller value occurs at large moments for continental events than for ridge and fracture zone events, suggesting systematic differences in stress drop.

(5) Other systematic factors affecting the calculation of M_s also appear to contribute to the observed regional bias.

We thank Professor J. H. Woodhouse for reading and correcting the manuscript and Professor H. Kanamori for constructive criticism throughout our work on this subject. This work was supported by the NSF.

Received 20 October 1987; accepted 4 February 1988.

1. Richter, C. F. *Bull. seism. Soc. Am.* **25**, 1–32 (1935).
2. Vanek, J. et al. *Izv. akad. Nauk. USSR, Ser. Geophys.* **2**, 153–158 (1962).
3. Aki, K. *Bull. Earthqu. Res. Inst. Tokyo Univ.* **44**, 23–88 (1966).
4. Agnew, D., Berger, J., Buland, R., Farrell, W. & Gilbert, F. *Eos* **57**, 280–288 (1976).
5. Peterson, J., Butler, H. M., Holcomb, L. G. & Hutt, C. R. *Bull. seism. Soc. Am.* **66**, 2049–2068 (1976).
6. Kanamori, H. & Given, J. W. *Phys. Earth planet. Inter.* **27**, 8–31 (1981).
7. Dziewonski, A. M., Chou, T. A. & Woodhouse, J. H. *J. geophys. Res.* **86**, 2825–2852 (1981).
8. Woodhouse, J. H. & Dziewonski, A. M. *J. geophys. Res.* **88**, 3247–3271 (1983).
9. Woodhouse, J. H. & Dziewonski, A. M. *J. geophys. Res.* **89**, 5953–5986 (1984).
10. Dziewonski, A. M., Franzen, J. E. & Woodhouse, J. H. *Phys. Earth planet. Inter.* **34**, 209–219 (1984).
11. Dziewonski, A. M., Ekström, G., Franzen, J. E. & Woodhouse, J. H. *Phys. Earth planet. Inter.* **45**, 11–36 (1987).
12. Dziewonski, A. M., Ekström, G., Woodhouse, J. H. & Zwart, G. *Phys. Earth planet. Inter.* (in the press).
13. Kanamori, H. *J. geophys. Res.* **82**, 2981–2987 (1977).
14. Richter, C. F. *Elementary Seismology* (W. H. Freeman, San Francisco, 1958).
15. Lienkamper, J. J. *Bull. seism. Soc. Am.* **74**, 2357–2378 (1984).
16. Kanamori, H. & Anderson, D. L. *Bull. seism. Soc. Am.* **65**, 1073–1095 (1975).
17. Ekström, G. & Dziewonski, A. M. *Bull. seism. Soc. Am.* **75**, 23–39 (1985).
18. Sipkin, S. A. *Bull. seism. Soc. Am.* **76**, 1515–1541 (1986).
19. Harkrider, D. G. *Bull. seism. Soc. Am.* **54**, 627–679 (1964).
20. Gutenberg, B. & Richter, C. F. *Gerlands Beitr. z. Geophysik* **47**, 73–131 (1936).
21. Gutenberg, B. *Bull. seism. Soc. Am.* **35**, 3–12 (1945).
22. Von Seggern, D. *Bull. seism. Soc. Am.* **60**, 503–516 (1970).
23. Nuttli, O. *Tectonophysics* **118**, 161–174 (1985).
24. Kanamori, H. & Allen, C. R. in *Maurice Ewing Series Vol. 6, Earthquake Source Mechanics* (American Geophysical Union, Washington, DC, 1986).
25. Zhuo, T. & Kanamori, H. *Bull. seism. Soc. Am.* **77**, 514–529 (1987).

Reshaping human antibodies for therapy

Lutz Riechmann[†], Michael Clark^{*}, Herman Waldmann^{*} & Greg Winter[‡]

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

^{*} Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK

A human IgG1 antibody has been reshaped for serotherapy in humans by introducing the six hypervariable regions from the heavy- and light-chain variable domains of a rat antibody directed against human lymphocytes. The reshaped human antibody is as effective as the rat antibody in complement and is more effective in cell-mediated lysis of human lymphocytes.

IN 1890 it was shown that resistance to diphtheria toxin could be transferred from one animal to another by the transfer of serum. It was concluded that the immune serum contained an anti-toxin, later called an antibody¹. For many years animal antisera were used in the treatment of microbial infections and for the neutralization of toxins in man². More recently rodent monoclonal antibodies (mAbs)³ have been used as 'magic bullets'⁴ to kill and to image tumours^{5,6}. The foreign immunoglobulin, however, can elicit an anti-globulin response which may interfere with therapy⁷ or cause allergic or immune complex hypersensitivity². Thus ideally human antibodies would be used. Human immunoglobulins are widely used as both prophylactic and microbicidal agents⁸, but it would be far better to have available human mAbs of the desired specificity. It has proven difficult, however, to make such mAbs by the conventional route of immortalization of human antibody-producing cells⁹.

There is an alternative approach. Antibody genes have been transfected into lymphoid cells, and the encoded antibodies expressed and secreted; by shuffling genomic exons, simple chimaeric antibodies with mouse variable regions and human constant regions have been made^{10–12}. Such chimaeric antibodies

have at least two advantages over mouse antibodies. First, the effector functions can be selected or tailored as desired. For example, of the human IgG isotypes, IgG1 and IgG3 appear to be the most effective for complement and cell-mediated lysis^{13–15}, and therefore for killing tumour cells. Second, the use of human rather than mouse isotypes should minimize the anti-globulin responses during therapy^{16,17} by avoiding anti-isotypic antibodies. The extent to which anti-idiotypic responses to rodent antibodies in therapy are dictated by foreign components of the variable versus the constant region is not known, but the use of human isotypes should reduce the anti-idiotypic response. For example, when mice were made tolerant to rat immunoglobulin constant-region determinants, administration of rat anti-lymphocyte antibodies did evoke anti-idiotypic responses, but these were delayed and weaker than in animals that had not been made tolerant¹⁸. Nevertheless, it is likely that a chimaeric antibody would provoke a greater immune response than a human mAb.

We have attempted to build rodent antigen binding sites directly into human antibodies by transplanting only the antigen binding site, rather than the entire variable domain, from a rodent antibody. The antigen binding site is essentially encoded by the hypervariable loops at one end of the β -sheet framework. The hypervariable regions of the heavy chain of mouse antibodies against a hapten¹⁹ or a protein antigen⁴⁷ were previously transplanted into a human heavy chain, and, in association with the mouse light chain, the antigen binding site was retained.

[†] Address from April 1988: Department of Molecular Biology, The Research Institute of Scripps Clinic, North Torrey Pines Road, La Jolla, California 02937, USA

[‡] To whom correspondence should be addressed.

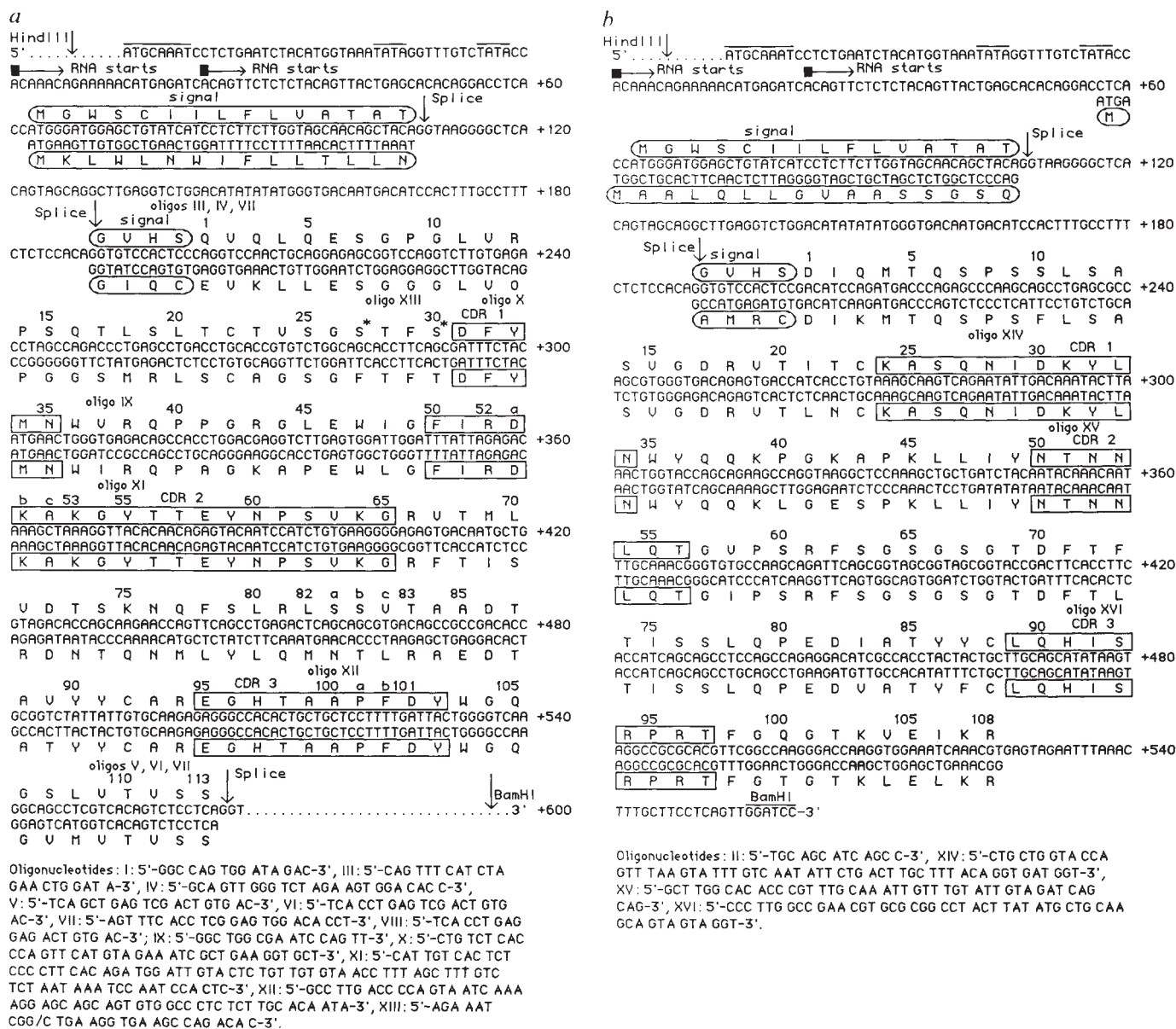


Fig. 1 Heavy-chain (a) and light-chain (b) sequences of the variable domains of reshaped (upper line) or rat YTH 34.5HL (lower line) antibodies. The reshaped heavy-chain variable domain HuVHCAMP was based on the HuVHNP gene^{12,19}, with the framework regions of human NEW (see note) alternating with the hypervariable regions of rat YTH 34.5HL. The reshaped light-chain variable domain HuVLCAMP is a similar construct, except with the framework regions of the human myeloma protein REI, with the C-terminal and the 3' non-coding sequence taken from a human J_κ-region sequence³⁶. The sequences of oligonucleotide primers are given and their locations on the genes are marked.

Methods. Messenger mRNA was purified³⁷ from the hybridoma clone YTH 34.5HL ($\gamma 2a$, κ^b). First strand cDNA was synthesized by priming with oligonucleotides complementary to the 5' end of the CH1 (oligonucleotide I) and the C κ exons (oligonucleotide II), and then cloned and sequenced as described previously^{38,39}. Two restriction sites (*Xba*I and *Sal*I) were introduced at each end of the rat heavy-chain variable region RaVHCAMP cDNA clone in M13 using mutagenic oligonucleotides III and V respectively, and the *Xba*I-*Sal*I fragment was excised. The corresponding sites were introduced into the M13-HuVHNP gene using oligonucleotides IV and VI, and the region between the sites was then exchanged. The sequence at the junctions was corrected with oligonucleotides VII and VIII, and an internal *Bam*HI site removed using the oligonucleotide IX, to create the M13-RaVHCAMP gene. The encoded sequence of the mature domain is thus identical to that of YTH 34.5HL. The reshaped heavy-chain variable domain (HuVHCAMP) was constructed in an M13 vector by priming with three long oligonucleotides simultaneously on the single strand containing the M13-HuVHNP gene^{12,19}. Each oligonucleotide (X, XI and XII) was designed to replace each of the hypervariable regions with the corresponding region from the heavy chain of the YTH 34.5HL antibody. Colony blots were probed initially with the oligonucleotide X and hybridization positives were sequenced: the overall yield of the triple mutant was 5%. The (Ser27 → Phe) and (Ser27 → Phe, Ser30 → Thr) mutants of M13mp8-HuVHCAMP were made with the mixed oligonucleotide XIII. The reshaped light-chain variable domain (HuVLCAMP) was constructed in M13 from a gene with framework regions based on human REI (J. Foote, unpublished data). As above, three long oligonucleotides (XIV, XV and XVI) were used to introduce the hypervariable regions of the YTH 34.5HL light chain.

Note: There are discrepancies involving the first framework region and the first hypervariable loop of the NEW heavy chain between the published sequence²⁷ used here and the sequence deposited in the Brookhaven data base (in parentheses): Ser27 (→ Thr), Thr28 (→ Ser) and Ser30 (→ Asp). Neither version is definitive (R. J. Poljak, personal communication) and the discrepancies do not affect our interpretations.

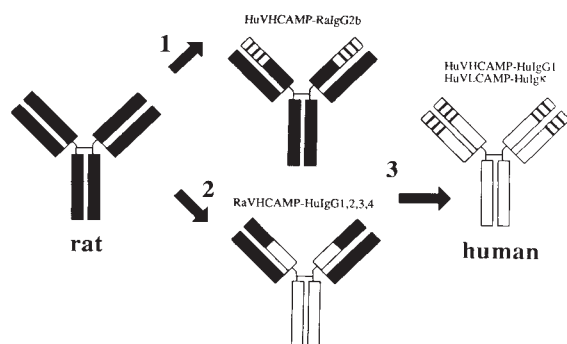


Fig. 2 Strategy for reshaping a human antibody for therapy. Sequences of rat origin are marked in black, and those of human origin in white. The recombinant heavy and light chains are also marked using a systematic nomenclature. See text for description of stages 1, 2 and 3. The genes encoding the variable domains were excised from the M13 vectors as *HindIII*-*Bam*HI fragments, and recloned into pSV2gpt²⁹ (heavy chains) or pSV2neo³⁰ (light chains), expression vectors containing the immunoglobulin enhancer¹². The human γ 1 (ref. 40), γ 2 (ref. 41), γ 3 (ref. 42), γ 4 (ref. 41) and κ (ref. 36) and the rat γ 2b (ref. 43) constant domains were introduced as *Bam*HI fragments. The following plasmids were constructed and transfected into lymphoid cell lines by electroporation⁴⁴. In stage 1, the pSVgpt plasmids HuVHCAMP-RaIgG2b, HuVHCAMP(Ser $\rightarrow$ Phe)-RaIgG2b, HuVHCAMP-(Ser27 $\rightarrow$ Phe, Ser30 $\rightarrow$ Thr)-RaIgG2b were introduced into the heavy chain loss variant of YTH 34.5HL. In stage 2, the pSVgpt plasmids RaVHCAMP-RaIgG2b, RaVHCAMP-HuIgG1, RaVHCAMP-HuIgG2, RaVHCAMP-HuIgG3, RaVHCAMP-HuIgG4 were transfected as above. In stage 3, the pSV-gpt plasmid Hu(Ser27 $\rightarrow$ Phe, Ser30 $\rightarrow$ Thr)VHCAMP-HuIgG1 was co-transfected with the pSV-neo plasmid HuVLCAMP-HuIgK into the rat myeloma cell line Y0 (Y B2/3.0 Ag 20 (ref. 31)). In each of the three stages, clones resistant to mycophenolic acid were selected and screened for antibody production by ELISA assays. Clones secreting antibody were subcloned by limiting dilution (for Y0) or the soft agar method (for the loss variant) and assayed again before 1 litre growth in roller bottles.

Since, to a first approximation, the sequences of hypervariable regions do not contain characteristic rodent or human motifs, such 'reshaped' antibodies should be indistinguishable in sequence from human antibodies.

There are mAbs to many cell-type-specific differentiation antigens, but only a few have therapeutic potential. Of particular interest is a group of rat mAbs directed against an antigen, the 'CAMPATH-1' antigen, which is strongly expressed on virtually all human lymphocytes and monocytes, but is absent from other blood cells including the haemopoietic stem cells²⁰. The CAMPATH-1 series contains rat mAb of IgM, IgG2a and IgG2c isotypes²¹, and more recently IgG1 and IgG2b isotypes which were isolated as class-switch variants from the IgG2a-secreting cell line YTH 34.5HL²². All of these antibodies, except the rat IgG2c isotype, are able to lyse human lymphocytes efficiently with human complement. Also the IgG2b antibody YTH 34.5HL-G2b, but not the other isotypes, is effective in antibody-dependent cell-mediated cytotoxicity (ADCC) with human effector cells²². These rat mAbs have important applications in problems of immunosuppression: for example control of graft-versus-host disease in bone-marrow transplantation²⁰; the management of organ rejection²³; the prevention of marrow rejection; and the treatment of various lymphoid malignancies (ref. 24 and M. J. Dyer, Hale, G., Hayhoe, F. G. J. and Waldmann, H., unpublished observations). The IgG2b antibody YTH 34.5HL-G2b seems to be the most effective at depleting lymphocytes *in vivo* but the use of all of these antibodies is limited by the anti-globulin response which can occur within two weeks of the initiation of treatment²⁴. Here we describe the reshaping of human heavy and light chains towards binding the

Table 1 Reshaping the heavy-chain variable domain

Heavy chain variable domain	Concentration of antibody in $\mu\text{g ml}^{-1}$ at	
	50% antigen binding	50% complement lysis
RaVHCAMP	0.7	2.1
HuVHCAMP	27.3	*
HuVHCAMP (Ser27 $\rightarrow$ Phe)	1.8	16.3
HuVHCAMP (Ser 27 $\rightarrow$ Phe, Ser 30 $\rightarrow$ Thr)	2.0	17.6

Antibodies with the heavy-chain variable domains listed above, rat IgG2b constant domains and rat light chains were collected from supernatants of cells at stationary phase and concentrated by precipitation with ammonium sulphate, followed by ion exchange chromatography on a Pharmacia MonoQ column. The yields of antibody were measured by an enzyme-linked immunosorbent assay (ELISA) directed against the rat IgG2b isotype, and each was adjusted to the same concentration³⁵. To measuring binding to antigen, partially purified CAMPATH-1 antigen was coated onto microtitre wells and bound antibody were detected via a biotin-labelled anti-rat IgG2b mAb³⁵, developed with a streptavidin-peroxidase conjugate (Amersham). Complement lysis of human lymphocytes was with human serum as the complement source²¹. For both binding and complement assays, antibody titres were determined by fitting the data to a sigmoid curve by at least squares iterative procedure²¹.

* Complement lysis with the HuVHCAMP variable domain was too weak for the estimation of lytic titre.

CAMPATH-1 antigen and the selection of human effector functions to match the lytic potential of the rat IgG2b isotype.

Strategy

The amino-acid sequences of the heavy- and light-chain variable domains of the rat IgG2a CAMPATH-1 antibody YTH 34.5HL were determined from the cloned complementary DNA (Fig. 1), and the hypervariable regions were identified according to Kabat²⁵. In the heavy-chain variable domain there is an unusual feature in the framework region. In most known heavy-chain sequences Pro41 and Leu45 are highly conserved: Pro41 helps turn a loop distant from the antigen binding site and Leu45 is in the β bulge which forms part of the conserved packing between heavy- and light-chain variable domains²⁶. In YTH 34.5HL these residues are replaced by Ala41 and Pro45 and presumably this could have some effect on the packing of the heavy- and light-chain variable domains. Working at the level of the gene and using three large mutagenic oligonucleotides for each variable domain, the rat hypervariable regions were mounted in a single step on the human heavy- or light-chain framework regions taken from the crystallographically solved proteins NEW²⁷ and REI²⁸ respectively (Fig. 1). The REI light chain was used because there is a deletion at the beginning of the third framework region in NEW. The reshaped human heavy- and light-chain variable domains were then assembled with constant domains in three stage (Fig. 2). This permits a step-wise check on the reshaping of the heavy-chain variable domain (stage 1), the selection of the human isotype (stage 2), and the reshaping of the light-chain variable domain and the assembly of human antibody (stage 3). The plasmid constructions were genomic, with the sequences encoding variable domains cloned as *HindIII*-*Bam*HI fragments and those encoding the constant domains as *Bam*HI-*Bam*HI fragments in either pSVgpt (heavy chain)²⁹ or pSVneo (light chain)³⁰ vectors. The heavy-chain enhancer sequence was included on the 5' side of the variable domain, and expression of both light and heavy chains was driven from the heavy-chain promoter and the heavy-chain signal sequence.

Heavy-chain variable domain

In stage 1, the reshaped heavy-chain variable domain (HuVHCAMP) was attached to constant domains of the rat

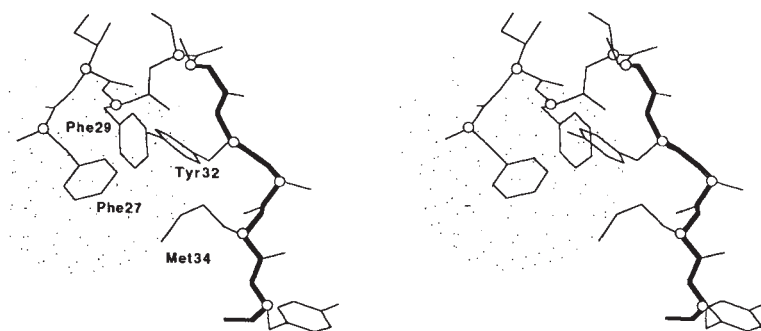


Fig. 3 Loop Phe27 to Tyr35 in the heavy-chain variable domain of the human myeloma protein KOL, which has been solved crystallographically⁴⁵. The backbone of the hypervariable region according to Kabat²⁵ is highlighted, and a 200% van der Waal surface is thrown around Phe 27 to show the interactions with Tyr 32 and Met 34 of the Kabat hypervariable region. In the rat YTH 34.5HL heavy chain, these three side chains are conserved in character, but in HuVHCAMP, Phe27 is replaced by Ser.

isotype IgG2b and transfected into a heavy-chain loss variant of the YTH 34.5 hybridoma. This variant carries two light chains, one derived from the Y3 fusion partner³¹. The cloned rat heavy-chain variable domain (RaVHCAMP) was also expressed as above, and the antibodies were purified and quantified (Table 1). The HuVHCAMP and RaVHCAMP antibodies, each of the rat IgG2b isotype, were compared to the CAMPATH-1 antigen in a direct binding assay and in complement lysis of human lymphocytes (Table 1). Compared with the original rat antibody, or the engineered equivalent, the antibody with the reshaped heavy-chain domain bound poorly to the CAMPATH-1 antigen and was weakly lytic. This suggested an error in the design of the reshaped domain.

There are several assumptions underlying the transfer of hypervariable loops from one antibody to another⁴⁷, in particular the assumption that the antigen binds mainly to the hypervariable regions. These are defined as regions of sequence²⁵ or structural³² hypervariability, the locations of hypervariable regions being similar by both criteria except for the first hypervariable loop of the heavy chain. By sequence the first hypervariable loop extends from residues 31–35 (ref. 25) whereas by structure it extends from residues 26–32 (ref. 32). Residues 29 and 30 form part of the surface loop, and residue 27, which is phenylalanine or tyrosine in most sequences, including YTH 35.5HL, helps pack against residues 32 and 34 (Fig. 3). Unlike most human heavy chains, in NEW (see note in Fig. 1) the phenylalanine is replaced by serine, which would be unable to pack in the same way. To restore the packing of the loop, we made both a Ser 27 → Phe mutation, and a Ser 27 → Phe, Ser 30 → Thr double mutation in HuVHCAMP. These two mutants showed a significant increase in binding to CAMPATH-1 antigen and lysed human lymphocytes with human complement (Table 1). Thus the affinity of the reshaped antibody could be restored by a single Ser 27 → Phe mutation, possibly as a consequence of an altered packing between the hypervariable regions and the framework. This suggests that alterations in the 'Kabat' framework region can enhance the affinity of the antibody and extends previous work in which an engineered change in the hypervariable region yielded an antibody with increased affinity³³.

Heavy-chain constant domains

In stage 2 (Fig. 2), the rat heavy-chain variable domain was attached to constant domains of the human isotypes IgG1, 2, 3 and 4, and transfected into the heavy-chain loss variant of the YTH 34.5 hybridoma. In complement lysis (Fig. 4a), the human IgG1 isotype proved similar to the YTH 34.5HL-G2b, with the human IgG3 isotype being less effective. The human IgG2 isotype was only weakly lytic and the IgG4 isotype was non-lytic. In ADCC (Fig. 4b) the human IgG1 was more lytic than the YTH 34.5HL-G2b antibody. The decrease in lysis at higher concentrations of the rat IgG2b and the human IgG1 antibody is due to an excess of antibody, which causes the lysis of effector cells. The human IgG3 antibody was weakly lytic, and the IgG2 and IgG4 isotypes were non-lytic.

We therefore selected the human IgG1 isotype for the reshaped antibody. Other recent work also favours the use of IgG1 isotype for therapeutic application. When the effector functions of human isotypes were compared using a set of chimaeric antibodies with an anti-hapten variable domain, the IgG1 isotype appeared superior to the IgG3 in both complement and cell-mediated lysis¹⁵. Also, of two mouse chimaeric antibodies with human IgG1 or IgG3 isotypes directed against cell surface antigens as tumour cell markers, only the IgG1 isotype mediated complement lysis^{13,14}.

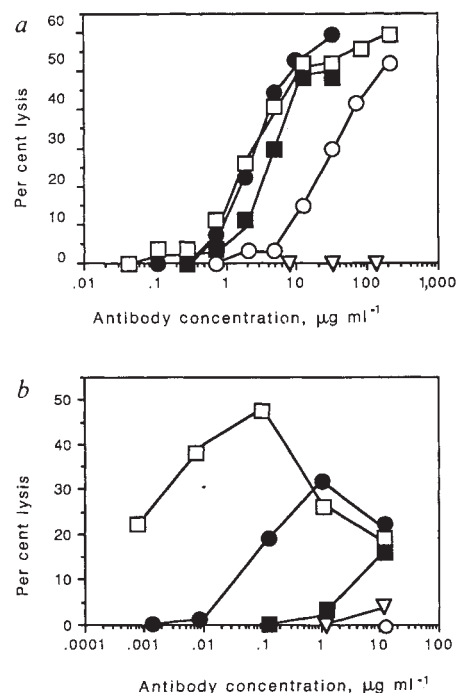


Fig. 4 *a*, Complement lysis and *b*, ADCC for antibodies with rat light-chain and rat heavy-chain variable domain attached to human IgG1 (□), IgG2 (○), IgG3 (■), or IgG4 (▽) isotypes. Lysis with the YTH 34.5HL antibody (●) is also shown.

Methods. Antibody was collected from cells in stationary phase, concentrated by precipitation with ammonium sulphate and desalted into phosphate buffered saline (PBS). Antibodies bound to the CAMPATH-1 antigen-coated on microtitre plates, were assayed in ELISA directed against the rat κ light chain³⁵, and each adjusted to the same concentration. The antibodies were assayed in complement lysis (Table 1) and ADCC with activated human peripheral blood mononuclear cells^{35,46}. Briefly, 5×10^4 human peripheral blood cells were labelled with ^{51}Cr and incubated for 30 min at room temperature with different concentrations of antibody. Excess antibody was removed and a 20-fold excess of activated cells added as effectors. After 4 h at 37 °C cell death was estimated by ^{51}Cr release.

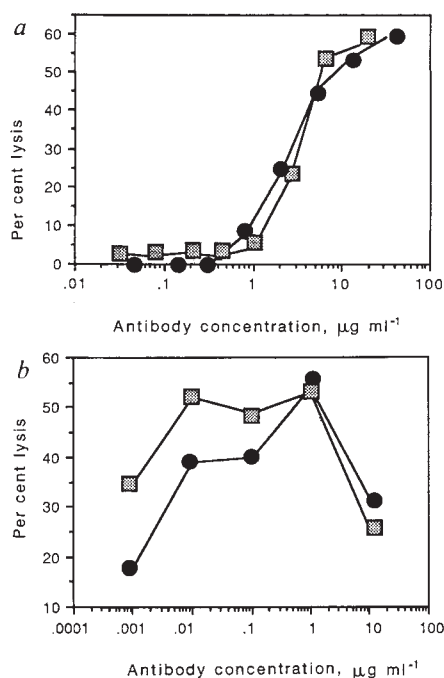


Fig. 5 *a*, Complement lysis and *b*, ADCC of the reshaped human (□) and rat YTH 34.5HL (●) antibodies. Antibody HuVHCAMP (Ser27→Phe, Thr30→Ser)–HuIGG1, HuVLCAMP–HuIGK was purified from supernatants of cells in stationary phase by affinity chromatography on protein-A Sepharose. The yield (about 10 mg l⁻¹) was measured spectrophotometrically. Complement and ADCC assays were performed as in Fig. 4.

Light chain

In stage 3 (Fig. 2), the reshaped heavy chain was completed by attaching the reshaped HuVHCAMP (Ser27→Phe, Ser30→Thr) domain to the human IgG1 isotype. The reshaped light-chain domain HuVLCAMP was attached to the human Cκ domain. The two clones were co-transfected into the non-secreting rat Y0 myeloma line. The resultant antibody, bound to CAMPATH-1 antigen (data not shown), and proved almost identical to the YTH 34.5HL-G2b antibody in complement lysis (Fig. 5*a*). In cell-mediated lysis the reshaped human antibody was more effective than the rat antibody (Fig. 5*b*). Similar results were

obtained with three different donors of target and effector cells (data not shown). Also, the antibody was as effective as YTH 34.5HL-G2b in killing leukaemic cells from three patients with B-cell lymphocytic leukaemia by complement-mediated lysis with human serum. Thus, by transplanting the hypervariable regions from a rodent to a human antibody of the IgG1 subtype, we have reshaped the antibody for therapeutic application.

Prospects

The availability of a reshaped human antibody with specificity for the CAMPATH-1 antigen should permit a full analysis of the *in vivo* potency and immunogenicity of an anti-lymphocyte antibody with wide therapeutic potential. Even if anti-idiotypic responses are eventually observed, considerable therapeutic benefit could be derived from an extended course of treatment. Also, it should be possible to circumvent an anti-globulin response restricted to idiotypic by using a series of antibodies with different idiotypes³⁴. In principle, the idiotypic of the reshaped CAMPATH-1 could be changed by altering the hypervariable regions or the framework regions—evidence from a reshaped antibody specific for the hapten nitrophenyl acetate suggests that recognition by anti-idiotypic antisera and anti-idiotypic mAbs is influenced by residues in the framework region¹⁹. Thus, recycling the hypervariable regions on different human framework regions should change the idiotypic, although ultimately it might focus the response directly onto the binding site for the CAMPATH-1 antigen. Although such focusing would be undesirable for CAMPATH-1 antibodies, it could be an advantage for the development of anti-idiotypic vaccines. It is likely that the answers to some of these questions will emerge from the use of this reshaped antibody in therapy.

We thank J. Foote for the synthetic gene encoding the framework regions of a human light chain; P. Leder, T. R. Rabbitts, T. Honjo, M. P. Lefranc respectively for clones encoding the constant regions of human κ chain, human IgG2 and IgG4, human IgG1, human IgG3; G. Hale for CAMPATH-1 antigen and for advice; M. Bruggemann and M. S. Neuberger for sub-clones of the heavy-chain constant regions¹⁵ and for advice; M. Frewin for technical assistance and C. Milstein for encouragement. L.R. is supported by a fellowship from the German 'Sonderprogramm Gentechnologie des DAAD'. The work was supported by the Medical Research Council and by Wellcome Biotechnology Ltd. 'CAMPATH' is a trademark of The Wellcome Foundation Ltd.

Received 3 December 1987; accepted 12 February 1988.

- Behring, E. von & Kitasato, S. *Dtsch. Med. Wochenschr.* **16**, 1113–1114 (1890).
- Ratner, B. *Allergy Anaphylaxis and Immunotherapy* (Williams and Wilkins, Baltimore, 1943).
- Kohler, G. & Milstein, C. *Nature* **256**, 495–497 (1975).
- Ehrlich, P. in *Immunology and Cancer Research* vol. 2 (ed. Himmelweit, F.) (Pergamon, London, 1957).
- Deland, F. H. & Goldenberg, D. M. in *Radionuclide Imaging* (ed. Kuhl, D. E.) 289–307 (Pergamon, Paris, 1982).
- Levy, R. & Miller, R. A. *Rev. Med.* **34**, 107–116 (1983).
- Miller, R. A., Oseroff, A. R., Stratte, P. T. & Levy, R. *Blood* **62**, 988–995 (1983).
- Leung, D. Y. M., Rosen, G. & Geha, R. S. in *Primary and Secondary Immunodeficiency Disorders* (ed. Chandra, R. K.) 44–61 (Churchill Livingstone, New York, 1983).
- Carson, D. A. & Freemark, B. D. *Adv. Immun.* **38**, 275–311 (1986).
- Morrison, S. L., Johnson, M. J., Herzenberg, S. A. & Oi, V. T. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6851–6855 (1984).
- Boulianne, G. L., Hozumi, N. & Shulman, M. J. *Nature* **312**, 643–646 (1984).
- Neuberger, M. S. *et al. Nature* **314**, 268–270 (1985).
- Liu, A. Y. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 3439–3443 (1987).
- Shaw, D. R. *et al. J. Immun.* **138**, 4534–4538 (1987).
- Bruggemann, M. *et al. J. exp. Med.* **166**, 1351–1361 (1987).
- Hale, G., Swirsky, D. M., Hayhoe, F. G. J. & Waldmann, H. *Molec. biol. Med.* **1**, 321–334 (1983).
- Giorgi, J. V. *et al. Transplant Proc.* **15**, 639–642 (1983).
- Benjamin, R., Cobbold, S. P., Clark, M. R. & Waldmann, H. *J. exp. Med.* **163**, 1539–1552 (1986).
- Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S. & Winter, G. *Nature* **321**, 522–525 (1986).
- Hale, G. *et al. Blood* **62**, 873–882 (1983).
- Hale, G., Hoang, T., Prospero, T., Watt, S. M. & Waldmann, H. *Molec. biol. Med.* **1**, 305–319 (1983).

- Hale, G. *et al. J. Immun. Meth.* **103**, 59–67 (1987).
- Hale, G., Waldmann, H., Friend, P. & Calne, R. *Transplantation* **42**, 308–311 (1986).
- Hale, G., Swirsky, D. M., Hayhoe, F. G. J. & Waldmann, H. *Molec. biol. Med.* **1**, 321–334 (1983).
- Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M. & Gottesman, K. S. in *Sequences of Proteins of Immunological Interest* (US Dept. of Health and Human Services, US Government Printing Office, 1987).
- Chothia, C., Novotny, J., Brucoleri, R. & Karplus, M. *J. molec. Biol.* **186**, 651–663 (1985).
- Saul, F. A., Amzel, M. & Poljak, R. J. *J. biol. Chem.* **253**, 585–597 (1978).
- Epp, O. *et al. Eur. J. Biochem.* **45**, 513–524 (1974).
- Mulligan, R. C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2072–2076 (1981).
- Southern, P. J. & Berg, P. *J. molec. appl. Genetics* **1**, 327–345 (1981).
- Galfre, G. & Milstein, C. *Meth. Enzym.* **73**, 1–46 (1981).
- Chothia, C. & Lesk, A. *J. Mol. Biol.* **196**, 901–917 (1987).
- Roberts, S., Cheetham, J. C. & Rees, A. R. *Nature* **328**, 731–734 (1987).
- Jonker, M. & den Brok, J. H. A. M. *Eur. J. Immun.* **17**, 1547–1552 (1987).
- Clark, M. & Waldmann, H. *J.N.C.I.* **79**, 1393–1398 (1987).
- Hieter, P. A., Max, E. E., Seidmann, J. G., Maizel, J. V. Jr & Leder, P. *Cell* **22**, 197–207 (1980).
- Kaartinen, M. *et al. J. Immun.* **130**, 320–324 (1983).
- Gubler, U. & Hoffmann, B. *J. Gene* **25**, 263–269 (1983).
- Sanger, F., Nicklen, S. A. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Takahashi, N., Ueda, N. S., Obata, M., Nikaido, T. & Honjo, T. *Cell* **29**, 671–679 (1982).
- Flanagan, J. G. & Rabbitts, T. H. *Nature* **300**, 709–713 (1982).
- Huck, S., Fort, P., Crawford, D. H., Lefranc, M.-P. & Lefranc, G. *Nucleic Acids Res.* **14**, 1779–1789 (1986).
- Bruggemann, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 6075–6079 (1986).
- Potter, H., Weir, L. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7161–7163 (1984).
- Marquardt, M., Deisenhofer, J., Huber, R. & Palm, W. *J. molec. Biol.* **141**, 369–391 (1980).
- Hale, G., Clark, M. & Waldmann, H. *J. Immun.* **123**, 3056–3061 (1985).
- Verhoeyen, M., Milstein, C. & Winter, G. *Science* (in the press).

Probes of large-scale structure in the Universe

Yasushi Suto, Krzysztof Górski*, Roman Juszkiewicz†
& Joseph Silk

Department of Astronomy, University of California, Berkeley,
California 94720, USA

Recent progress in observational techniques has made it possible to confront quantitatively various models for the large-scale structure of the Universe with detailed observational data. In particular, there are well-established galaxy–galaxy and cluster–cluster correlations^{1–3}, stringent upper limits on (and one tentative measurement of) the anisotropy of the cosmic microwave background radiation (CBR) on various angular scales^{4,5}, and large-scale bulk motions reported by several groups^{6–8}. These three observables place important constraints on cosmological parameters, most notably the amplitude and shape of the density fluctuation spectrum in the very early Universe. We develop a general formalism to show that the gravitational instability theory for the origin of large-scale structure is now capable of critically confronting observational results on CBR angular anisotropies, large-scale bulk motions and large-scale clumpiness in the galaxy counts.

On large scales, where linear perturbation theory is valid, the mass excess $(\delta M/M)^2$ averaged over regions with characteristic size r_M , the peculiar velocity v averaged over r_v , and the correlation function $C(\theta)$ of intrinsic CBR temperature fluctuations on angular scale θ due to the Sachs–Wolfe effect are expressed as^{9–11}

$$\left(\frac{\delta M}{M}\right)^2(r_M) = \frac{v^2}{2\pi^2} \int_0^\infty k^2 P(k) W_M(kr_M) dk \quad (1)$$

$$v^2(r_v) = \frac{\Omega^{1.2} H^2}{2\pi^2} \int_0^\infty P(k) W_v(kr_v) dk \quad (2)$$

and

$$C(\theta) = \frac{\Omega^{0.6} H^4}{8\pi^2} \int_0^\infty k^{-2} P(k) \frac{\sin R_H \theta k}{R_H \theta k} W_T(kr_T) dk \quad (3)$$

where $R_H = 2c/H\Omega$, c is the velocity of light, $P(k)$, Ω and H respectively denote the power spectrum of the irregularities, the cosmological density parameter, and the Hubble constant evaluated at the present time. The function $W_T(kr_T)$, approximately given by $\exp(-k^2 r_T^2)$, represents the effect of finite thickness of the last scattering surface (for instantaneous recombination, $r_T = 0$). As one does not know whether the luminous matter in the Universe is a reliable indicator of the underlying matter distribution, the notion of biasing has been introduced to model this uncertainty, and the parameter ν in equation (1) allows us to incorporate the biasing effect. The quantities v^2 and $C(\theta)$ in equations (2) and (3) respond directly to the gravitational field, and are independent of whether light traces mass or not. For $\Omega < 1$ equation (3) is approximate and assumes $\theta \ll \Omega$ radians. For $\Omega = 1$, the expression for $C(\theta)$ is exact⁹.

The window functions W_M and W_v describe how the data on positions and velocities of galaxies (and clusters) are smoothed to extract the large-scale behaviour of these observables. If this smoothing is represented as a convolution of density or velocity fields with some smoothing function $w(r)$ (normalized to unity and assumed to be bounded), then $W(k) = \hat{w}(k)^2$, where $\hat{w}(k)$ is the three-dimensional Fourier transform of $w(r)$. Provided one considers slowly varying power spectra $P(k)$, as in a cold

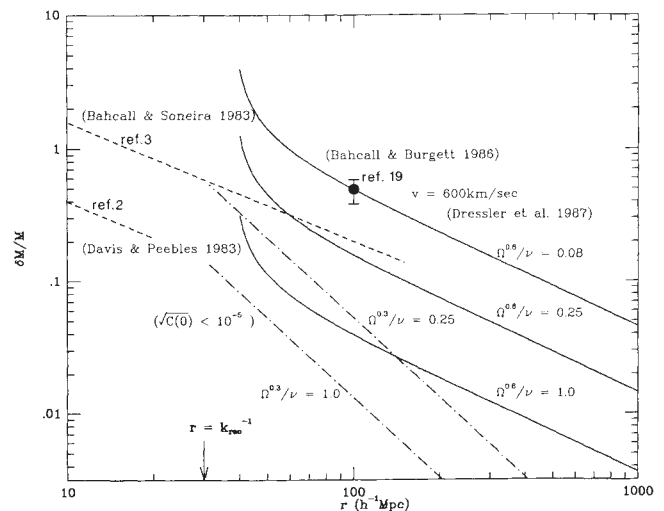


Fig. 1 Upper limits on $(\delta M/M)(r)$ derived from relations (4) and (5), using $v = 600 \text{ km s}^{-1}$ at $r_v = 40 h^{-1} \text{ Mpc}$. We plot the cases for $\Omega^{0.6}/\nu = 1.0, 0.25$ and 0.08 in solid lines. The dash-dotted lines come from equation (5) by assuming that the upper limit on any intrinsic $\sqrt{C(0)}$ (equation 3) is 10^{-5} for $\Omega^{0.3}/\nu = 0.25$ and 1.0 . For comparison, the expected amplitudes of $\delta N/N$ given by equation (13) are plotted using the galaxy–galaxy², cluster–cluster³ and supercluster–supercluster¹⁹ correlation functions (dashed lines and filled circle).

dark matter model, for example, the particular choice of the smoothing function $w(r)$ is not important. In the following analysis, however, we shall try to develop predictions of CBR temperature fluctuations $\delta T/T$ that are independent of the form of $P(k)$ and which should be valid in general models for large-scale structure. We therefore consider the properties of smoothing functions more carefully. (One can in fact prove relations similar to equations (4)–(6) below using smoothing prescriptions for which $\hat{w}(k)$ is always positive, but this constraint does not translate clearly to the required properties of $w(r)$. However, we are able to present analytical examples that are qualitative but quite illustrative.)

Consider the following two functions and their Fourier transforms:

$$w_1(r) \propto (1+r^2)^{-n}, \quad \hat{w}_1(k) \propto e^{-k} \sum_{m=0}^{n-2} \frac{(2n-m-4)! (2k)^m}{(n-m-2)! m!}$$

and

$$w_2(r) \propto (1+r^{2n})^{-1},$$

$$\hat{w}_2(k) \propto \sum_{m=1}^n \exp(-k \sin((m-\frac{1}{2})\pi/n)) / k$$

$$\times \cos[(m-\frac{1}{2})2\pi/n + k \cos((m-\frac{1}{2})\pi/n)]$$

where $n > 2$. While $w_1(r)$ and $w_2(r)$ have the same asymptotic behaviour for $r \rightarrow \infty$, $\hat{w}_1(k)$ is always positive, but $\hat{w}_2(k)$ changes sign. This critical difference results from the behaviour of the smoothing functions around $r=1$; $w_1(r)$ decreases rather gradually, while $w_2(r)$ drops rapidly, approaching a step-function as n becomes large. Clearly, any rapid variations in the smoothing function enhance the small-scale contribution to averages such as in equations (1) and (2). This was already emphasized^{10,12} for the particular example of step-function (or top-hat) smoothing. Our example of $w_1(r)$ shows that there exists a (presumably) broad class of functions with $\hat{w}(k) > 0$, which are also good indicators of large scale properties of density and/or velocity fields. A convenient choice, which we will subsequently use, is the gaussian smoothing function, for which

* Permanent address: Copernicus Astronomical Centre, Poland.

† Present address: Department of Physics, Princeton University, Princeton, New Jersey 08544, USA; permanent address: Warsaw University Observatory, Poland.

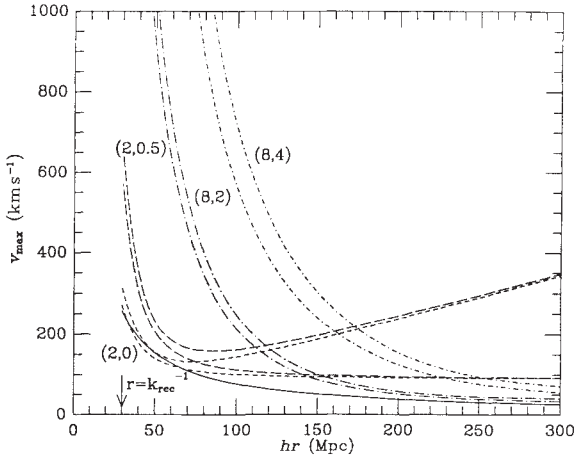


Fig. 2 Upper limits on $v(r)$ derived from relation (6). Using the power spectrum which minimizes C , given in equation (3), we calculate the observable $\delta T(\theta, \sigma)/T$ on angular scale θ with antenna beam-size σ . The curves are labelled by (θ, σ) . The dashed and the dot-dashed lines correspond to observations on 2° and 8° , respectively. For each parameter, we plot the predictions for a triple beam-switching (upper-curve) and a double beam-switching (lower curve) experiment. It is interesting to note that given the upper limit on $\delta T/T$, a double beam-switching sets more stringent constraints on $v(r)$ on larger scales. All curves are plotted assuming that the r.m.s. value of $\delta T/T$ is less than 10^{-5} . Thus in the case of upper limits at the 95% confidence level, for example, $\delta T/T$ should be considered to be less than 2×10^{-5} . The solid curve corresponds to the case of an 'ideal' experiment with double beam switching ($\sigma = 0$, and θ is large enough to ensure that $\delta T(\sigma, \theta)/T$ is given by $\sqrt{2C(0)}$). In this case, we plot the curve by assuming that $\delta T/T = \sqrt{2C(0)} < 10^{-5}$.

$W(k) = \exp(-k^2)$. A gaussian window function is commonly used and has the desirable properties of damping out fluctuations due to small-scale noise in the data sample¹²⁻¹⁵.

Applying the mean value theorem⁹ to equations (1)–(3), and setting $\theta = 0$, we find the following inequalities:

$$v^2(r_v) \geq \frac{e\Omega^{1.2}H^2}{\nu^2} (r_M^2 - r_v^2) \left(\frac{\delta M}{M}\right)^2 (r_M) \quad (4)$$

$$C(0) \geq \frac{e^2\Omega^{0.6}H^4}{16\nu^2} (r_M^2 - r_T^2) \left(\frac{\delta M}{M}\right)^2 (r_M) \quad (5)$$

$$C(0) \geq \frac{e\Omega^{-0.6}H^2}{4} (r_v^2 - r_T^2) v^2(r_v) \quad (6)$$

where $e = 2,718$. The above argument works only when the relation $r_M > r_v > r_T$ is satisfied. The equalities in the above expressions hold when $P(k) \propto \delta(k - 1/\sqrt{r_M^2 - r_v^2})$, $\delta(k - 1/\sqrt{r_M^2 - r_T^2})$ and $\delta(k - 1/\sqrt{r_v^2 - r_T^2})$, respectively, where $\delta(x)$ is the Dirac delta function. The crucial assumption here is that a purely gravitational instability is responsible for the formation of large-scale structure in the Universe. The above inequalities then hold for any particular form of $P(k)$.

Another interesting relation among these quantities is obtained by applying the Cauchy-Schwarz inequality to equations (1) to (3). For gaussian window functions, it reduces to

$$\left(\frac{\delta M}{M}\right)^2 (r_M) C(0) \geq \frac{\nu^2\Omega^{-1.8}}{4} v^4 \left(\sqrt{\frac{r_M^2 + r_T^2}{2}}\right) \quad (7)$$

This relation, in which the scales r_M and r_T are arbitrary, is important in considering the amplitudes of $\delta M/M$ and v on the same scales where (4) is not useful. Using relations (4) to (7) together with appropriate observational data, one can

attempt to obtain various constraints on the behaviour of $(\delta M/M)^2(r_M)$, $v^2(r_v)$ and $C(0)$.

Before presenting our results, several remarks on the validity of relations (4)–(7) are useful. First, our present analysis is based on linear perturbation theory. Roughly speaking, the coherence length of the galaxy-galaxy correlation function $r_c \approx 8h^{-1}$ Mpc gives a measure of the nonlinear scale (here we use the dimensionless Hubble constant $h = H/(100 \text{ km s}^{-1} \text{ Mpc}^{-1})$). Therefore our formalism is valid for $r_M > r_v \gg r_c$. This condition is well satisfied on those scales in which we are interested (see Figs 1 and 2 below). Non-standard theories such as cosmic strings¹⁶ and explosive amplification scenarios^{17,18}, which involve either intrinsically non-gaussian fluctuations, nonlinear objects or nongravitational forces, are beyond the scope of the present analysis. The length scale associated with the fuzziness of the last scattering surface⁹ is small compared with r_v and r_M , provided that $r_v, r_M \gg r_c$. Thus we can safely set $r_T = 0$ in what follows.

Second, we consider angular scales where the Sachs-Wolfe term is a dominant source of the CBR anisotropy and we neglect the possibility of reionization after recombination. Other contributions to the CBR anisotropies, especially due to the Doppler term, can be neglected for scales larger than the horizon size at recombination, $k_{\text{rec}}^{-1} \approx 30\Omega^{-1/2}h^{-1}$ Mpc (ref. 9). However, because this scale is larger than the observed scales over which possible streaming motions are seen when Ω is small, our relations (5) and (6) should be considered with caution if $\Omega < 1$.

Finally, it should be noted that we underestimate the actual $C(0)$ for isocurvature perturbations where the phase of primordial $\delta T/T$ on the last scattering surface is the same as in the Sachs-Wolfe term¹⁰; in this case, the constraints are expected to be stronger than in the adiabatic case.

Rough estimates of inequalities (4) to (7) yield:

$$\left(\frac{\delta M}{M}\right) \leq 4 \times 10^{-2} \nu \Omega^{-0.6} \left(\frac{100h^{-1} \text{ Mpc}}{r_M}\right) \left(\frac{v}{600 \text{ km s}^{-1}}\right) \quad (r_M \gg r_v) \quad (8)$$

$$\left(\frac{\delta M}{M}\right) \leq 1.3 \times 10^{-2} \nu \Omega^{-0.3} \left(\frac{100h^{-1} \text{ Mpc}}{r_M}\right)^2 \left(\frac{\sqrt{C}}{10^{-5}}\right) \quad (r_M \gg r_T) \quad (9)$$

$$v \leq 200\Omega^{0.3} \left(\frac{60h^{-1} \text{ Mpc}}{r_v}\right) \left(\frac{\sqrt{C}}{10^{-5}}\right) \text{ km s}^{-1} \quad (r_v \gg r_T) \quad (10)$$

$$v \leq 1,000\nu^{-1}\Omega^{0.9} \left(\frac{\delta M}{M}\right) \left(\frac{\sqrt{C}}{6 \times 10^{-6}}\right) \text{ km s}^{-1} \quad \left(r_v = \sqrt{\frac{r_M^2 + r_T^2}{2}}\right) \quad (11)$$

Interestingly, the ranges of the amplitudes of $(\delta M/M)$, v and $\sqrt{C}$ that appear in the above inequalities are very close to the observed values. Note that relation (8) is not affected by possible reionization after recombination and that relation (10) is independent of any biasing.

The intrinsic anisotropy $C(0)$ is not directly observable. To make a more realistic comparison between theory and experiment, we simulate the actual CBR anisotropy observations on a specific scale θ with an antenna beam-width σ by convolving $C(\theta)$ with a gaussian beam of dispersion σ (ref. 9). Our strategy here is to find the observable $\delta T(\theta, \sigma)/T$ under the condition of minimizing the total intrinsic anisotropy $C(0)$: we are not trying to minimize $\delta T/T$ itself.

The upper limits on $\delta M/M$ as a function of scale r which are expected from the observed bulk motions⁸ are plotted in Fig. 1 (because we use a gaussian window function, we rescale the claimed value of $r_v = 60h^{-1}$ Mpc to $40h^{-1}$ Mpc; for further discussion of the observations, see ref. 20). The uncertainties associated with the estimate of r_v do not change the present results for $r_M \gg r_v$ (equation 9). For comparison, we also calculate the amplitude of r.m.s. fluctuations in the galaxy number counts $\delta N/N$ from the observed correlation function $\xi(x)$. We

use a gaussian window function according to equation (1), and obtain

$$\left(\frac{\delta N}{N}\right)^2(r) = \frac{1}{2\sqrt{\pi}r^3} \int_0^\infty x^2 \xi(x) \exp\left(-\frac{x^2}{4r^2}\right) dx \quad (12)$$

Since the integrand in equation (12) has a gaussian cutoff on large scales, the final $\delta N/N$ is insensitive to the asymptotic behaviour of $\xi(x)$. Thus we assume $\xi(x) \propto x^{-1.8}$ for $x > 0$ simply by extrapolating the power-law tail to obtain

$$\left(\frac{\delta N}{N}\right)(r) \approx 0.7\sqrt{\xi(r)} \quad (13)$$

which is plotted in Fig. 1 (dashed lines).

Inspection of Fig. 1 suggests that clusters of galaxies and superclusters appear to have formed in preferentially high density locations ($\nu \gg 1$). Forthcoming CBR observations on angular scales of several degrees should set stringent constraints on the amplitude of peculiar velocity and mass irregularities over large scales. The dash-dotted line in Fig. 1 is the maximum $\delta M/M$ assuming that the total intrinsic $\delta T/T = \sqrt{C}$ is less than 10^{-5} . Because of the r^{-2} dependence in equation (9), an upper limit on $\delta T/T$ with this sensitivity, if available, would impose a stringent constraint on $\delta M/M$ over large scales. In Fig. 2, we use a $P(k)$ which minimizes $C(0)$ to derive the maximum amplitude of the bulk motion as a function of r . Figure 2 is based on relation (6), assuming that the resultant $\delta(\theta, \sigma)/T$ should be less than 10^{-5} . It is clear that an observation over 2° with $\sim 0.5^\circ$ antenna beam-size would be of particular interest in testing the presence of the claimed large bulk motions⁷⁻⁹.

The treatment described here is based on several approximations. None of them is expected to change the present results significantly, although any inferences should be made with caution. Also it is important to note that the inequalities (4)–(7) should be regarded as rather conservative, because we cannot by any means expect the ideal delta function to represent $P(k)$ in a realistic astrophysical situation. Hence we expect that our conclusions should in general be even more stringent, despite the approximations used here. This already suggests that all presently advocated cosmological models will have considerable difficulty in simultaneously explaining the observed amplitudes (1)–(3), should the observations indeed be confirmed. Finally we emphasize that the search for anisotropy in the CBR on angular scales of a few degrees is of great value in probing the origin of large-scale structure in the Universe (Fig. 2). It is remarkable that with the present observational results (or with those available in the near future), we can address, without resorting to any specific model, the issue of whether or not gravitational instability accounts for the formation of large-scale structure.

We thank M. Davis for stimulating discussions and R. F. G. Wyse for a critical reading of the manuscript. Y.S. gratefully acknowledges the support of the Miller Institute for Basic Research in Science at the University of California, Berkeley. This research has also been supported in part by grants from NSF and NASA.

Received 29 September; accepted 16 December 1987.

1. Groth, E. J. & Peebles, P. J. E. *Astrophys. J.* **217**, 385–405 (1977).
2. Davis, M. & Peebles, P. J. E. *Astrophys. J.* **267**, 465–482 (1983).
3. Bahcall, N. A. & Soneira, R. M. *Astrophys. J.* **270**, 20–38 (1983).
4. Kaiser, N. & Silk, J. *Nature* **324**, 529–537 (1986).
5. Davies, R. D. *et al. Nature* **326**, 462–465 (1987).
6. Rubin, V. C. *et al. Astr. J.* **81**, 687–737 (1976).
7. Collins, C. A., Joseph, R. D. & Robertson, N. A. *Nature* **320**, 506–508 (1986).
8. Dressler, A. *et al. Astrophys. J.* **313**, L37–L42 (1987).
9. Juszkiewicz, R., Górski, K. & Silk, J. *Astrophys. J.* **323**, L1–L6 (1987).
10. Peebles, P. J. E. *The Large Scale Structure of the Universe* (Princeton Univ. Press, 1980).
11. Sachs, R. W. & Wolfe, A. N. *Astrophys. J.* **147**, 73–90 (1967).
12. Zel'dovich, Ya. B. & Novikov, I. D. *Relativistic Astrophysics. II: The Structure and Evolution of the Universe* (Univ. Chicago Press, 1983) section 13.3.
13. Vittorio, N. & Silk, J. *Astrophys. J.* **293**, L1–L4 (1985).
14. Vittorio, N., Juszkiewicz, R. & Davis, M. *Nature* **323**, 132–133 (1986).
15. Vittorio, N. & Turner, M. S. *Astrophys. J.* **316**, 475–482 (1987).

16. Silk, J. & Vilenkin, A. *Phys. Rev. Lett.* **53**, 1700–1703 (1984).
17. Ikeuchi, S. *Publ. astr. Soc. Japan* **33**, 211–222 (1981).
18. Ostriker, J. P. & Cowie, L. L. *Astrophys. J.* **243**, L127–L131 (1981).
19. Bahcall, N. A. & Burgett, W. S. *Astrophys. J.* **300**, L35–L39 (1986).
20. Bardeen, J. M., Bond, J. R. & Efstathiou, G. *Astrophys. J.* **321**, 28–35 (1987).

Neutron star/red giant encounters in globular clusters

Charles D. Bailyn

Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA

It has been suggested that the tidal capture of a red giant by a neutron star in globular clusters may result in bound systems¹, the subsequent evolution of which may lead to a variety of interesting objects, including millisecond pulsars^{2,3}. But, the binding energy of the envelope of a giant is much less than that of a main-sequence star⁴, so the ratio $x_{\text{crit}} = R_{\text{crit}}/R_*$ (where R_{crit} is the maximum distance of closest approach between two stars for which the tidal energy is sufficient to bind the system, and R_* is the radius of the star on which tides are being raised) will be less for giants than for main-sequence stars. Here I present a simple expression for the amount by which x_{crit} is diminished as a star evolves, and conclude that tidal capture of giants by neutron stars resulting in binary systems is unlikely in globular clusters. However, collisions between neutron stars and red giants^{3,5,6}, or an alternative process involving tidal capture of a main-sequence star into an initially detached binary system⁷, may result either in rapidly rotating neutron stars or in white dwarf/neutron star binaries.

The tidal oblateness ε of a star of mass M_* caused by a nearby star of mass M is approximately $\varepsilon \approx f M x^{-3}/M_*$ where $x = d/R_*$, d is the distance between the two stars, and f is a constant ≤ 1 depending on the frequency associated with the encounter and the frequency of the induced oscillations⁸. The energy associated with such a tidal bulge is $E \approx (GM_* M_e/R_*)\varepsilon^2$, where M_e is the mass of the star outside a point mass core. For a tidal capture to occur, this energy must be greater than the total energy in the orbital motion of the two stars. When the two stars are far from one another, this energy can be written as $E_{\text{orb}} \approx \mu v_d^2/2$ where v_d is the velocity dispersion of the cluster, and the reduced mass $\mu = M_* M/(M_* + M)$. Therefore a tidal capture requires that at periastron

$$x \leq \left[f^2 \frac{GM_*}{R_{\text{ms}} v_d^2} \frac{M(M+M_*)}{M_*^2} \right]^{1/6} \left[\frac{M_e}{M_*} \frac{R_{\text{ms}}}{R_*} \right]^{1/6} = x_{\text{crit}} \quad (1)$$

where R_{ms} is the radius of a main-sequence star of mass M_* .

The first term in brackets in equation (1) is the result obtained for a main-sequence star⁸, while the second term represents the ratio between x_{crit} for tidal capture of a main-sequence star and of a giant of the same mass, assuming that f is the same in both cases. If we assume a value of $x_{\text{crit}} = 3$ for main-sequence stars^{9–11}, we can calculate x_{crit} for any subsequent evolutionary stage by determining the term in the second bracket from an appropriate sequence of stellar models. The results for a star of $M = 0.7 M_\odot$, with helium fraction $Y = 0.3$ and metal abundance $Z = 0.0001^{12,13}$ are presented in Fig. 1.

Figure 1a shows the value of x_{crit} as the star evolves from the zero-age main-sequence (ZAMS) to the base of the giant branch. There are only minor changes in x_{crit} until core hydrogen exhaustion, when the formation of a degenerate core and the rapid expansion of the star's envelope cause a rapid decrease in x_{crit} . More precise calculations¹¹ suggest a somewhat greater decrease in x_{crit} as the star evolves across the main-sequence, resulting from the decreasing half mass radius which is not considered

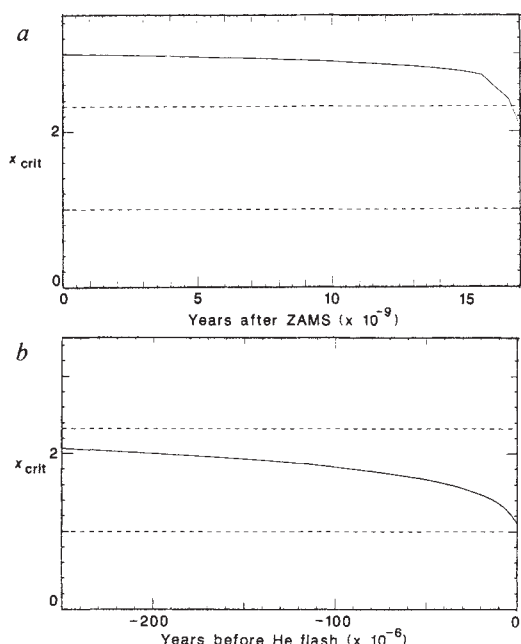


Fig. 1 Equation (1) applied to a $0.7M_{\odot}$ star with $Y=0.3$ and $Z=0.0001$. Encounters in which the closest approach is closer than the upper dotted line ($x \equiv R/R_{\odot} = 2.32$, where $R_{\odot}$ is the stellar radius) result in material being pulled off the giant by the neutron star. Encounters closer than the lower dotted line ($x = 1$) result in collisions.

here. The sharp decrease in x_{crit} after the star leaves the main-sequence continues as the star evolves up the giant branch, as shown in Fig. 1b. This decrease is again caused both by the decreasing binding energy of the envelope as the star expands, and by the increasing size of the 'point mass' degenerate core. By the time the star reaches the helium flash, x_{crit} drops to 1.1. If we assume that collisions ($x_{\text{crit}} \leq 1$) lead to some kind of coalesced object rather than to a binary system, the cross-section for the formation of neutron star/giant binaries when the giant is near the tip of the giant branch drops to 1/20 of the value obtained when x_{crit} is assumed to be equal to 3.

Taking into account the reduction of x_{crit} as the star evolves, one finds that 2% of all tidal captures will be of giants and subgiants, rather than 7% which is the value obtained when $x_{\text{crit}} = 3$ is assumed for all stars¹. Of these 2%, 3/4 will be captures of subgiants rather than giants, as opposed to 3/7 in the $x_{\text{crit}} = 3$ case. Thus the total number of tidally captured giants is reduced by a factor of 8 compared to previous estimates.

There are several additional reasons why even these numbers may seriously overestimate the number of bound giant/neutron star systems formed by tidal capture. First, encounters that are near collisions may have catastrophic results. The gravitational attraction of the passing neutron star on gas at the surface of the giant will be greater than that of the giant itself if $(x-1)^2 \leq M_n/M_g$ where x is the distance of closest approach in units of the giant's radius, and M_n and M_g are the masses of the neutron star and the giant respectively. For $M_n = 1.4M_{\odot}$ and $M_g = 0.8M_{\odot}$ (the turnoff mass of a galactic globular clusters), this critical closest approach will be at $x = 2.32$. Encounters closer than this will result in at least some material being pulled off the giant. It should be noted that as the orbit is nearly parabolic, the two stars will be near periastron only for a short amount of time. For this reason, the giant may lose only its outermost layers to the neutron star. Nevertheless, an encounter with a distance of closest approach only slightly larger than the radius of the giant, as would be required for a tidal capture of a star on the giant branch, is likely to be disastrous for the giant, especially as the close encounter will be repeated in subsequent orbits as the

system circularizes. Also, the energy deposited in tides will result in a considerable increase in the stellar radius, making the difficulties outlined here even more severe¹¹. Furthermore, the lifetimes of any bound systems which do form will be relatively short, since a captured giant or subgiant will quickly overflow its Roche lobe, and transfer mass on its short evolutionary timescale. Thus we might expect not to see any examples of giants tidally captured by neutron stars into bound systems. Detailed hydrodynamical simulations will be necessary to show whether neutron star/giant encounters form bound systems of any variety, and what kind of remnant is left orbiting the neutron star.

Fortunately, it is not necessary to postulate the formation of a neutron star/giant binary system to account for the isolated millisecond pulsar in M28, as collisions between red giants and neutron stars may lead to a similar evolutionary endpoint^{3,5}. Such collisions are likely to have occurred in at least some galactic globular clusters, and are likely to result in the neutron star spiralling into the envelope of the giant. The fate of such common envelope systems depends on the energy transport in the envelope. If enough energy is carried to the surface (by convection or other energy transport mechanisms) then the envelope may remain hydrostatic, and the spiralling-in process will continue until the neutron star and the degenerate core of the giant have coalesced. Otherwise, the energy deposited by dynamical friction in the envelope will eventually result in its expulsion¹⁴. If this expulsion occurs before the neutron star and the giant's degenerate core have spiralled together, but after they have approached close enough that gravitational radiation will bring them into contact in a Hubble time, a mass-transferring white dwarf/neutron star binary like 4U1820-30 (ref. 15) may result⁶.

A statistically more probable origin for 4U1820-30 has been proposed⁷, which might also account for the formation of rapidly rotating single neutron stars. In this scheme, a main-sequence star is tidally captured by a neutron star forming an initially detached binary system. The secondary fills its Roche lobe as it evolves off the main-sequence, resulting in unstable mass transfer⁷ and the formation of a common envelope. Like the giant collision scenario, such a system can then either evolve into a white dwarf/neutron star binary or a single rapidly rotating neutron star, depending on the energy transport processes in the common envelope. In ref. 7 it is assumed that $x_{\text{crit}} = 3.3$ as the star evolves from the ZAMS to the base of the giant branch, in which case tidal captures by neutron stars during this time are ~ 6 times as likely to occur as subsequent collisions with neutron stars⁷. However, the value $x_{\text{crit}} = 3.3$ is only valid for ZAMS stars with $M = 0.4M_{\odot}$, whereas a star which has subsequently evolved off the main-sequence would have $M \approx 0.8M_{\odot}$. Assuming $x_{\text{crit}} = 3.0$ for a ZAMS star, the calculated reduction of x_{crit} as the star evolves, and including collisions with all stars which have evolved off the main-sequence (both subgiants and giants), this ratio of probabilities is found to be ~ 2.5 .

The fate of white dwarf/neutron star binaries formed by neutron star/giant collisions may be influenced by the large mass transfer rate at the time when white dwarf and neutron stars initially come into contact⁷. If the core mass $M_{\text{core}} > 0.4M_{\odot}$, unstable mass transfer will occur¹⁶. If the mass transfer rate is larger than the Eddington limit (as expected if $M_{\text{core}} > 0.1M_{\odot}$) the necessary mass loss from the system may carry off enough angular momentum to spiral the white dwarf and the neutron star together on a dynamical timescale. In these cases a stable mass-transferring white dwarf/neutron star system like 4U1820-30 will not result. Instead a neutron star surrounded by a massive disk may arise, providing yet another route to the formation of a millisecond pulsar in a globular cluster.

I thank J. Grindlay for discussions and encouragement. This work was partially supported by the NSF. The author is a member of the Harvard Society of Fellows.

Received 26 October 1987; accepted 5 January 1988.

- Verbunt, F. & Hut, P. in *IAU Symp.* 125 (eds Helfand, D. J. & Huang, J. H.) 187–196 (Reidel, Boston, 1987).
- Romani, R. W., Kulkarni, S. R. & Blandford, R. D. *Nature* **329**, 309–310 (1987).
- Verbunt, F., van den Heuvel, E. P. J., van Paradijs, J. & Rappaport, S. A. *Nature* **329**, 312–314 (1987).
- Ostriker, J. P., in *IAU Symp.* 113 (eds Goodman, J. & Hut, P.) 358 (Reidel, Boston, 1985).
- Suntanyo, W. *Astr. Astrophys.* **44**, 227–230 (1975).
- Verbunt, F. *Astrophys. J.* **312**, L23–L25 (1987).
- Bailyn, C. D. & Grindlay, J. E. *Astrophys. J.* **316**, L25–L29 (1987).
- Fabian, A., Pringle, J. & Rees, M. J. *Mon. Not. R. Astr. Soc.* **172**, P15–P18 (1975).
- Press, W. H. & Teukolsky, S. A. *Astrophys. J.* **213**, 183–192 (1977).
- Lee, H. M. & Ostriker, J. P. *Astrophys. J.* **310**, 176–188 (1986).
- McMillan, S. L. W., McDermott, P. N. & Taam, R. E. *Astrophys. J.* **318**, 261–277 (1987).
- Mengel, J. G., Sweigart, A. V., Demarque, P. & Gross, P. G. *Astrophys. J. Suppl. ser.* **40**, 733–791 (1979).
- Sweigart, A. V. & Gross, P. G. *Astrophys. J. Suppl. ser.* **36**, 405–437 (1978).
- Bodenheimer, P. & Taam, R. E. in *NATO Conf: Evolution of Galactic X-ray Binaries* (eds Trumper, J., Lewin, W. H. G. & Brinkmann, W.) 13–24 (NATO ASI Series, 1986).
- Stella, L., Friedhorsky, W. & White, N. E. *Astrophys. J.* **312**, L17–22 (1987).
- Bonsema, P. T. J. & van den Heuvel, E. P. J. *Astr. Astrophys.* **146**, L3–L4 (1984).

Molecular resolution images of amino acid crystals with the atomic force microscope

Scot Gould*, O. Marti*, Barney Drake*,
Louis Hellemans*†, Charles E. Bracker*‡,
Paul K. Hansma*, Nancy L. Keder§, Michael M. Eddy§
& Galen D. Stucky§

* Department of Physics, University of California, Santa Barbara, California 93106, USA

† Department of Chemistry, University of Leuven, B-3030, Heverlee, Belgium

‡ Department of Botany and Plant Pathology, Purdue University, West Lafayette, Indiana 47907, USA

§ Department of Chemistry, University of California, Santa Barbara, California 93106, USA

The atomic force microscope has been used to image arrays of molecules at the surface of DL-leucine crystals. Lattice spacings are consistent with X-ray diffraction data. In contrast to metals and semiconductors, the surface of these amino acid crystals seems to be a simple termination of the bulk; there is no evidence of a surface reconstruction for this molecular crystal. This initial success in imaging amino acid molecules points to the potential usefulness of atomic force microscopy for imaging molecules of biological importance.

In the two years since its invention, the atomic force microscope (AFM)¹ has already been developed into a powerful analytical instrument. It has been used to obtain atomic resolution images of graphite^{2–4} and other layered and non-layered materials^{5,6}, to image the molecules in an organic monolayer⁷, to study details of friction with atomic resolution⁴, to characterize electronic materials⁸, and to image magnetic materials—even the recording head of a magnetic device⁹. It operates on principles similar to its predecessor, the scanning tunnelling microscope (STM)¹⁰, by mapping the contours of surfaces, but whereas the STM requires a specimen that will conduct electrons, the AFM can be used with specimens that are either insulating or conducting.

In essence the atomic force microscope works like a record player. A tip is pressed against the surface of a specimen with a minute tracking force of $\sim 10^{-8}$ N, and scanned across it. The images are composed of a series of scans, each displaced from the previous one by a small distance, typically 0.01 nm for the images shown here. Imaging a sample at these tiny tracking forces does not seem to damage the surface, as the same surface can be scanned repeatedly without changing the image.

The heart of an atomic force microscope is the force-sensing assembly. In all atomic force microscopes that have been built

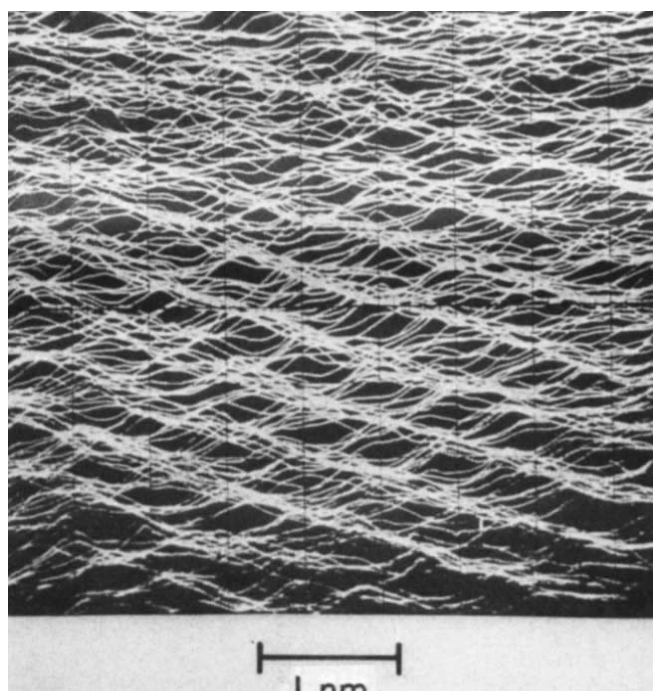


Fig. 1 Line scan AFM image of a crystal of DL-leucine (a photograph of raw data from an oscilloscope screen). The mounds are due to the tip of the atomic force microscope scanning over individual molecules of leucine.

to date^{1–9}, the force has been sensed by measuring the deflection of a small spring with a spring constant in the range 0.1–100 N m⁻¹. For the images in this paper, the spring was a cantilever constructed of gold-plated tungsten wire with a diameter of 20 μ m. The cantilever was in the shape of a 'V'; each side was 1.4 mm long, and a pointed fragment from a shattered diamond was mounted at the apex of the 'V' to act as the probing tip that contacted the surface of the sample. The spring constant was ~ 6 N m⁻¹.

The deflection of the cantilever was measured by detecting the electron-tunnelling current between the tungsten wire and a platinum-covered razor blade that had been positioned with mechanical and piezoelectric elements a small distance behind the cantilever³. For the images in this report the razor blade was positioned so that the deflection of the cantilever was ~ 1 nm. The forces were thus $\sim 6 \times 10^{-9}$ N. All the images presented are of the varying tunnelling current between the razor blade and the spring (variable deflection mode). This is comparable to the constant height mode used with scanning tunnelling microscopes. The slow feedback loop adjusted the sample position such that the average tunnelling current (and hence the average deflection of the cantilever) was kept constant.

For sample preparation, crystalline flakes of DL-leucine (Calbiochem no. 4320) were adhered to pieces ($\sim 5 \times 5$ mm) of glass microscope slide with a small amount of gap-filling cyanoacrylate glue and stored in a desiccator when not being used. The exposed face of the crystals was imaged directly in air with the AFM.

Figure 1 is a line scan AFM image of the DL-leucine in which each molecule appears as a small mound. About eight molecules are visible in each row. The spacing between the mounds is ~ 0.5 nm. Figure 2 shows a grey-scale image of a DL-leucine sample taken on another atomic force microscope in our laboratory. In this image each molecule is represented as a light patch against a darker background.

An X-ray diffraction pattern of the sample as prepared for AFM studies was recorded on a Scintag θ - θ diffractometer (Fig. 3). DL-leucine crystallizes^{11,12} in the triclinic space group *P1*, with lattice parameters $a = 0.539$, $b = 1.412$, $c = 0.519$ nm, $\alpha =$

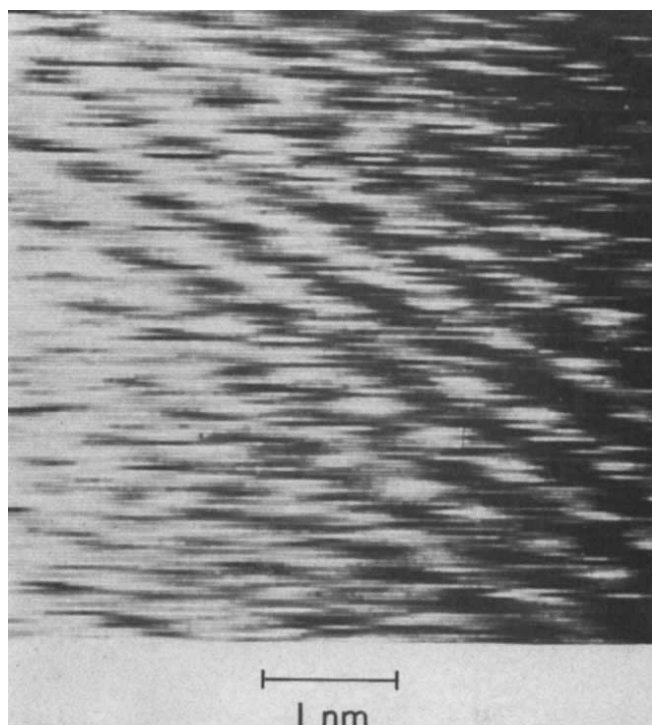


Fig. 2 Gray-scale image of DL-leucine taken on another AFM (a photograph of raw data from a video screen fed by an Arlunya Temporal Filter and Image Storage Device; Arlunya TF 5000 Image Processor, Princeton Electronic Products). The lattice of leucine molecules appears as lighter spots on a darker background.

97.0° , $\beta = 111.1^\circ$, $\gamma = 86.4^\circ$. The diffraction pattern shows a series of peaks at 2θ intervals of 6.26° ($\text{CuK}\alpha$, $\lambda = 0.15418 \text{ nm}$) corresponding to the $0k0$ reflections. Thus, the face of the crystal being imaged is parallel to the a - c plane of the unit cell.

The molecular graphics package CHEMX was used to generate the packed crystalline structure from the published crystallographic data^{11,12}. Examination of the structure parallel to the a - c plane indicates that the most likely point of cleavage is between the methyl groups of leucine (Fig. 4), rather than at the amino-acid functional groups at the other end of the molecules, because the van der Waals forces between the layers of non-polar groups in the crystal are much weaker than the hydrogen bonding that holds the double layer of carboxyl and amino groups together.

We can estimate some of the unit cell parameters from our AFM images, but sources of error must be taken into consideration. Thermal drift, insufficient lateral stiffness of the cantilever, sloped surfaces, and the calibration error of the piezos contribute to the error in measurement. Sloped surfaces reduce the measured distances below the true values, whereas the insufficient lateral stiffness increases the measured distances above the true values. Thermal drift can affect measurements in both directions with typical thermal drift rates of 0.01 – 0.1 nm s^{-1} . The horizontal (x) scanning speed used ($\sim 470 \text{ nm s}^{-1}$) was much larger than the thermal drift, so horizontal measurements were scarcely affected by drift. The vertical (y) scanning speed, on the other hand, was of the order of 0.5 nm s^{-1} , the same magnitude as the drift. Consequently, vertical measurements are less reliable. For example, for the image presented in Fig. 1, the measured y dimension is 12% less than the X-ray value^{11,12}. To estimate distances and their errors we include the calibration error for x and y (10%) and a drift error for y (20%). The image in Fig. 1 then gives axes of $0.54 \pm 0.06 \text{ nm}$ and $0.45 \pm 0.13 \text{ nm}$; Fig. 2 gives $0.58 \pm 0.09 \text{ nm}$ and $0.52 \pm 0.12 \text{ nm}$.

The measured angles are affected by the same factors as length

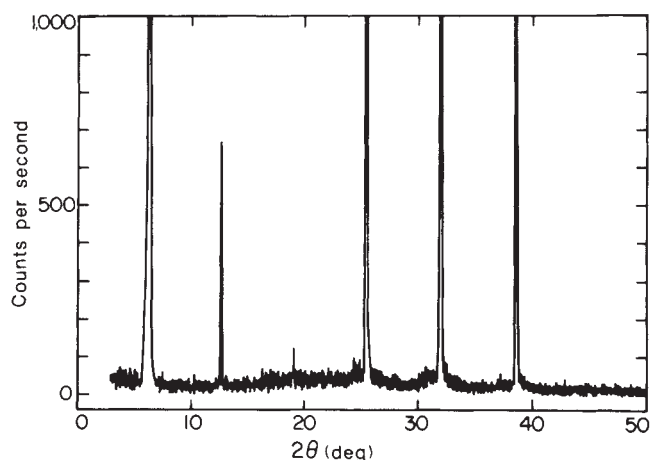


Fig. 3 X-ray diffraction pattern at a wavelength of 0.154 nm of DL-leucine crystals as mounted for the AFM scans. All peaks are consistent with the molecular orientation shown in Fig. 4. No peaks corresponding to any other orientation were detected.

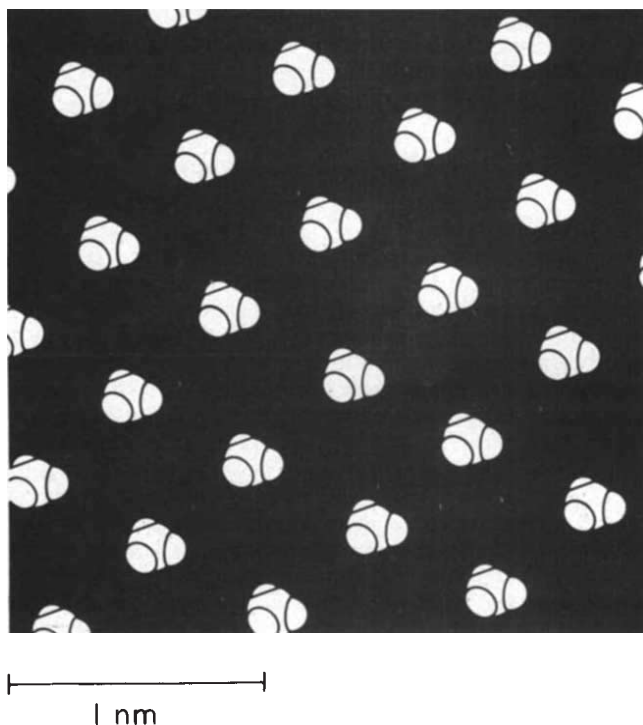
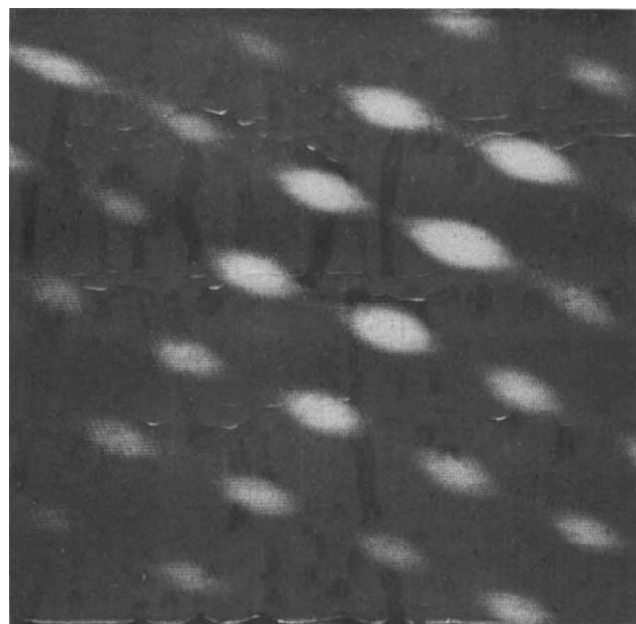


Fig. 4 Array of methyl groups spaced according to X-ray data (ref. 12), and rotated to an orientation similar to that in Fig. 5. Drawn with the program CHEMX (Chemical Design Ltd, Oxford, UK).

measurements, except for the calibration error which is equal for both scan axes. But error in the angle between the two scan axes (5°) has to be included. For example, for the image presented in Fig. 1 the angle is $110^\circ \pm 9^\circ$, whereas in Fig. 2 the angle is $120^\circ \pm 15^\circ$, both within range of the X-ray result of 111° (refs 11, 12).

Figure 5 is a Wiener filtered version¹³ of the data shown in Fig. 2 for comparison with the surface inferred from the X-ray diffraction results (Fig. 4). The agreement between Fig. 4, which is based on data dominated by molecules in the interior of DL-leucine crystals, and Fig. 5, which shows only molecules on the surface, demonstrates that the surface is a simple termination of the bulk; no surface reconstruction can be detected within the resolution of these experiments.



1 nm

Fig. 5 Computer-processed image of the raw data presented in Fig. 2b. The Wiener filtering decreased the noise without affecting the periodic data.

These are the first images of amino acids to come from the emerging STM/AFM technology, and they reflect early progress in our effort to explore the fine structure of biological materials. Despite some promising results¹⁴⁻¹⁸, it is still too soon to say how useful the scanning tunnelling microscope will be to the understanding of biological structure, mainly because the STM relies on electrical conductivity of the specimen to provide information. Biomolecules, for the most part, are not good conductors. The AFM, however, operates independently of specimen conductivity and may be better suited in the long run as a tool for studying atomic and molecular architecture of biological interest.

Although it has been useful for this initial research to obtain images from a periodic structure, it should be emphasized that periodic structures are not necessary for the atomic force microscope. In this respect the atomic force microscope is similar to a scanning tunnelling microscope. Many examples of isolated atomic scale features and point defects in periodic structures have been reported for the scanning tunnelling microscope¹⁹. It thus appears possible, at least in principle, to image molecules of biological importance that cannot be crystallized.

We thank W. Stoeckenius for emphasizing the opportunities for imaging proteins with the atomic force microscope, and I. Giaever for suggesting that amino acid crystals would be a good place to begin. We thank G. Binnig, Ch. Gerber, C. Quate, T. Albrecht, K. Wickramasinghe, R. Erlandsson, G. McClelland, C. Mate and S. Chiang for sharing advice on how to build atomic force microscopes and sending us preprints of their work; L. Dubois and M. Van Hove for information about surface reconstructions; S. Bartnicki-Garcia for providing the DL-leucine crystals; E. Martzen for secretarial help and M. Wilson and R. Stuber for help in the design and construction of our microscopes. This work was supported in part by the Belgian National Fund for Scientific Research (L.H.), the Agricultural Experiment Station, Purdue University Journal Paper No. 11, 597 (C.E.B.), Office of Naval Research (O.M., B.D.), and the National Science Foundation (S.G., P.K.H.).

Received 8 February; accepted 25 February 1988.

1. Binnig, G., Quate, C. F. & Gerber, Ch. *Phys. Rev. Lett.* **56**, 930-933 (1986).
2. Binnig, G., Gerber, Ch., Stoll, E., Albrecht, T. R. & Quate, C. F. *Europhys. Lett.* **3**, 1281-1286 (1987).
3. Marti, O., Drake, B. & Hansma, P. K. *Appl. Phys. Lett.* **51**, 484-486 (1987).
4. Mate, C. M., McClelland, G. M., Erlandsson, R. & Chiang, S. *Phys. Rev. Lett.* **59**, 1942-1945 (1987).
5. Albrecht, T. R. & Quate, C. F. *J. appl. Phys.* **62**, 2599-2602 (1987).
6. Marti, O., Drake, B., Gould, S. & Hansma, P. K. *J. Vac. Sci. Technol.* (in the press).
7. Marti, O. *et al. Science* **239**, 50-52 (1988).
8. Martin, Y., Williams, C. C. & Wickramasinghe, H. K. *J. Appl. Phys.* **61**, 4723-4729 (1987).
9. Martin, Y., Williams, C. C. & Wickramasinghe, H. K. *Appl. Phys. Lett.* **50**, 1455-1457 (1987).
10. Binnig, G. & Rohrer, H. *Rev. Mod. Phys.* **59**, 615-625 (1987).
11. *Crystal Data: Organic Compounds 1975-1978* 3rd edn Vol 5 (eds Donnay, J. D. H. & Ondik, H. M.) (National Bureau of Standards, Washington), 1983.
12. di Blasio, B., Pedone, C. & Sirigu, A. *Acta Crystallogr.* **B131**, 601, 2941 (1975).
13. Stoll, E. & Marti, O. *Surf. Sci.* **181**, 222-229 (1987).
14. Smith, D. P. E. *et al. Proc. natn. Acad. Sci.* **84**, 969-972 (1987).
15. Travaglini, G., Rohrer, H., Amrein, M. & Gross, H. *Surf. Sci.* **181**, 380-390 (1987).
16. Baro, A. M. *et al. Nature* **315**, 253-254 (1985).
17. Zasadinski, J. A. N., Schneir, J., Gurley, J., Elings, V. & Hansma, P. K. *Science* **239**, 1013-1015 (1988).
18. Lindsay, S. M. & Barris, B. J. *Vac. Sci. Technol. A* (in the press).
19. Slough, C. G., McNairy, W. W., Coleman, R. V., Drake, B. & Hansma, P. K. *Phys. Rev. B* **34**, 994-1005 (1986).

Crystallography, chemistry and structural disorder in the new high- T_c Bi-Ca-Sr-Cu-O superconductor

D. R. Veblen*, P. J. Heaney*, R. J. Angel†, L. W. Finger†, R. M. Hazen†, C. T. Prewitt†, N. L. Ross†, C. W. Chu‡, P. H. Hor‡ & R. L. Meng‡

* Department of Earth and Planetary Sciences, The Johns Hopkins University, Baltimore, Maryland 21218, USA
† Geophysical Laboratory, Carnegie Institution of Washington, Washington, DC 20008-3898, USA

‡ Department of Physics, Texas Center for Superconductivity and Space Vacuum Epitaxy Center, University of Houston, Houston, Texas 77004, USA

High-temperature superconductivity with onset temperatures up to ~ 120 K has been observed recently in the Bi-Ca-Sr-Cu-O system^{1,2}. The primary high-temperature superconducting phase has been identified by Hazen *et al.*² as a layer structure, probably related to the 22-K superconducting phase recently described by Michel *et al.*³ and to the family of layered bismuth compounds described by Aurivillius⁴. Hazen *et al.* described the unit cell, approximate composition, and electrical properties of the Bi-Ca-Sr-Cu-O superconducting phase. Diffraction experiments indicate that the new Bi-Ca-Sr-Cu-O layer-structure superconductor possesses a primitive orthorhombic unit cell with probable space group $Pnnn$. The material exhibits severe structural disorder, which is primarily related to stacking within the layers. The apparent orthorhombic structure is an average resulting from orthorhombic material mixed with monoclinic domains in two twinned orientations. We also describe two distinct types of structural disorder that are common in materials synthesized to date. This disorder complicates the crystallographic analysis and suggests that X-ray and neutron diffraction methods may yield only an average structure.

The present observations are from samples BCSCO-a and BCSCO-b of Hazen *et al.*². All selected-area electron diffraction (SAED), X-ray analytical electron microscopy (AEM), and high-resolution transmission electron microscopy (HRTEM) experiments were performed with a Philips 420 transmission electron microscope, as described in detail by Livi and Veblen⁵. Cliff and Lorimer⁶ thin-film k -factors for Ca and Cu were determined as noted in ref. 5, and those for Sr and Bi were determined from celestite (natural SrSO_4) and synthetic BiVO_4 , combined with previously determined k -factors. The analyses are not highly accurate but they are clearly from single-crystal

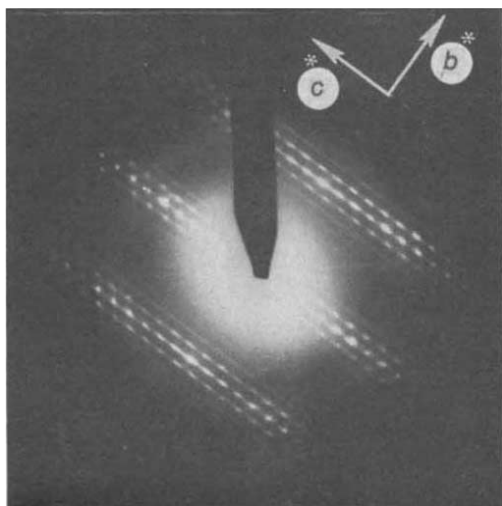


Fig. 1 An a -axis SAED pattern showing intense streaking parallel to c^* , due to severe structural disorder.

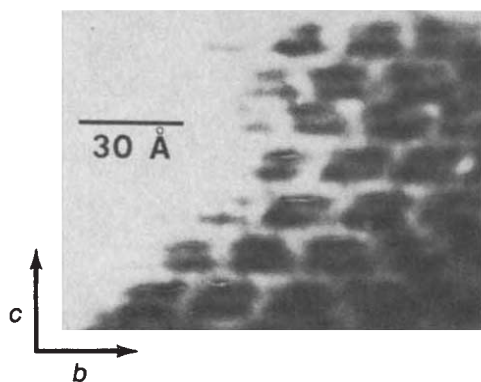


Fig. 3 Thin-crystal HRTEM image obtained parallel to a and labelled with the ideal orthorhombic unit cell axes. Light fringes parallel to (001) show the basic layer structure, while the light '010' fringes exhibit variable orientation as a result of stacking disorder.

material of the superconducting phase and not from a mixture. The quality of the lattice images is rather low, because the layered crystals are quite thick in orientations with the electron beam parallel to the layers and because most crystals are deformed and hence do not represent a unique crystallographic orientation. They are, however, of sufficient quality to allow determination of the large-scale defect structures.

SAED patterns obtained parallel to the a and c crystallographic axes suggest² an orthorhombic unit cell. This identification was refined using powder X-ray diffraction data. The unit cell has parameters $a = 5.41$, $b = 27.2$, $c = 30.8$ Å, which may be related to a cubic perovskite subcell by the ratios $\sqrt{2} \times 5\sqrt{2} \times \sim 8$. Based on the systematic absence of diffraction spots with $h = 0$, $k + l = 2n + 1$, on a -axis SAED patterns, Hazen *et al.*² suggested that the oxide might possess an A -centred unit cell.

We have now obtained both SAED patterns and HRTEM lattice images parallel to all three crystallographic axes as well as other orientations. These patterns confirm that the unit cell is orthogonal, with all three unit cell angles $\alpha = \beta = \gamma = 90^\circ$, consistent with orthorhombic lattice geometry. Intensity distributions also are consistent with the orthorhombic crystal system.

In addition to the systematic absence noted above, b -axis SAED patterns show that diffraction spots with $k = 0$, $h + l = 2n + 1$, are absent, consistent with a body-centred (I -centred) lattice or n -glide planes parallel to (010). The a - and b -axis SAED patterns indicate that most apparent violations of I -

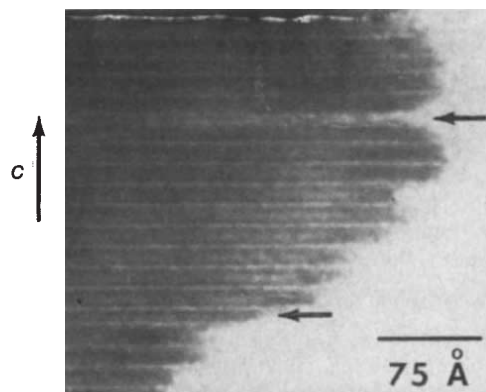


Fig. 2 Lattice image showing the 15.5-Å (002) fringes. Two defects involving missing or extra sublayers are arrowed.

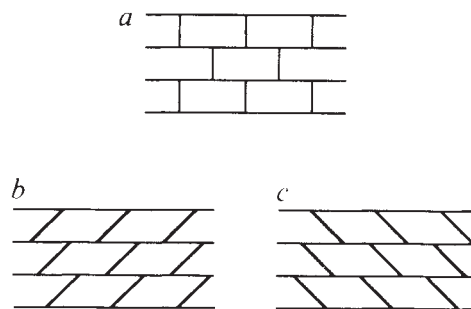


Fig. 4 Schematic diagram of the structure observed in lattice images from thin crystals; a , ideal orthorhombic structure with vertical '010' fringes; b , and c , the two twin orientations of the monoclinic domains with tilted '010' fringes, which combine to form an average orthorhombic structure in electron and X-ray diffraction patterns.

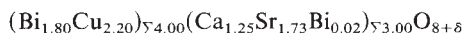
centring visible on the c -axis patterns result from the intersection with the Ewald sphere of relatively strong streaks that are present in all diffraction rows parallel to c^* , as a result of structural disorder. The intensities of the strongest apparent violations (the 100, 090, and 0-11-0 spots in Fig. 3 of Hazen *et al.*²) are probably enhanced by dynamical diffraction, due to their relationships to the strong perovskite subcell diffractions.

The electron diffraction extinctions noted above contain clear violations of face centring, as well as all possible orthorhombic end-centred lattices (F , A , B and C), and these violations cannot be accounted for by dynamical diffraction. In addition, diffractions such as the $15l$ with $l = 2n + 1$ are present, in violation of body-centring. The lattice thus appears to be primitive (P). The observed space group extinctions indicate n -glide planes normal to c , b and probably a , suggesting space group $Pnnn$, or possibly $Pnn2$ or $Pnnm$. It should further be noted that entire classes of reflections, such as the $15l$ s with $l = 2n$ are absent. These non-space-group extinctions probably imply that the structure possesses some sort of strong pseudosymmetry or local symmetry.

It must be emphasized that the unit cell and space-group symmetry under discussion refer only to the average structure of this superconducting oxide. It is clear from data presented below that the true structure consists of small domains probably having both monoclinic and orthorhombic symmetries.

AEM data from thin, single-crystal plates of the oxide show that the material is chemically rather homogeneous. As the AEM

analyses are from indisputably single-phase material that exhibited SAED patterns characteristic of the superconductor, these data support the assumption that the superconductor possesses a unique structural formula, although compositional variations among synthesis runs can be expected to occur due to solid solution. The AEM data show rough consistency with the data of Hazen *et al.*², although the latter used experimental techniques with much lower spatial resolution. When normalized to 7.00 cations, for comparison with the data of Hazen *et al.*, the analyses yield the formula



Determination of the full crystal structure is necessary for full understanding of the chemical substitutions within this formula.

SAED patterns of the oxide show reasonable periodicity parallel to a^* and b^* but tend to be heavily streaked parallel to c^* (Fig. 1), indicative of structural disorder affecting the d_{001} periodicity. Lattice images confirm that the structure is highly periodic within the a - b plane, but at least two different types of disorder are readily observed in HRTEM lattice images obtained by imaging diffraction patterns containing c^* . One-dimensional lattice images produced from only the 001 row of diffraction spots show occasional layers with contrast or spacing different from that of the bulk structure (Fig. 2; also see Fig. 4 in ref. 2). Assuming that the oxide is related to the perovskite structure, these defects are probably due to missing and extra perovskite sublayers, perhaps accompanied by anomalous cation occupancies. The most common defects of this sort appear to involve a single missing or extra unit of ~ 4 Å thickness parallel to the c -axis.

Another type of disorder, observed in images obtained with the electron beam parallel to a , appears to be pervasive and probably results in most of the streaking parallel to c^* in SAED patterns. These a -axis images (Fig. 3) show relatively light fringes that are rigorously parallel to (001) and another set of light fringes that are more or less parallel to (010), hereafter referred to as the '010' fringes. Although the average crystal structure of this oxide has not yet been determined, it is likely that these light (001) and '010' fringes correspond to two different sets of planes with relatively low electron and proton density.

Although the '010' fringes in lattice images are in places parallel to (010) of the orthorhombic unit cell, they commonly deviate from this orientation in either of the two possible senses, forming variable angles other than 90° with the (001) fringes. This is shown schematically in Fig. 4. The structures of Fig. 4b and c correspond to monoclinic domains related to each other by a (010) mirror twin operation. The true structure of this material may thus be a mixture of an orthorhombic structure (Fig. 4a) with twinned monoclinic structure (Fig. 4b and c), exhibiting variable angles between the (001) and (010) planes. Microtwinning of such monoclinic domains is shown in Fig. 5a. In addition, the mixing of these orthorhombic and monoclinic domains results in orientation variations in the 011 fringes on lattice images (Fig. 5b).

Although there is no direct diffraction data relating to the space group of the monoclinic structures, it is likely that it is a distortional subgroup of the ideal orthorhombic space group $Pnmm$. Assuming that the n -glide normal to a is preserved, the space group is probably $P2_1/n$, having a non-conventional monoclinic unit cell with a unique.

A final defect that is observed in a -axis lattice images can be described as a dislocation in the superstructure with apparent projected Burgers vector $\frac{1}{2}[011]$. An example is illustrated in Fig. 5c. This type of defect can be accommodated by adjusting the stacking of (001) layers in the region of the fault.

We have clarified the crystallography of the new high- T_c Bi-Ca-Sr-Cu-O layered superconductor, showing that its average structure is probably best described with a primitive orthorhombic unit cell, with space group $Pnmm$. However, our lattice imaging experiments also show that the material grown

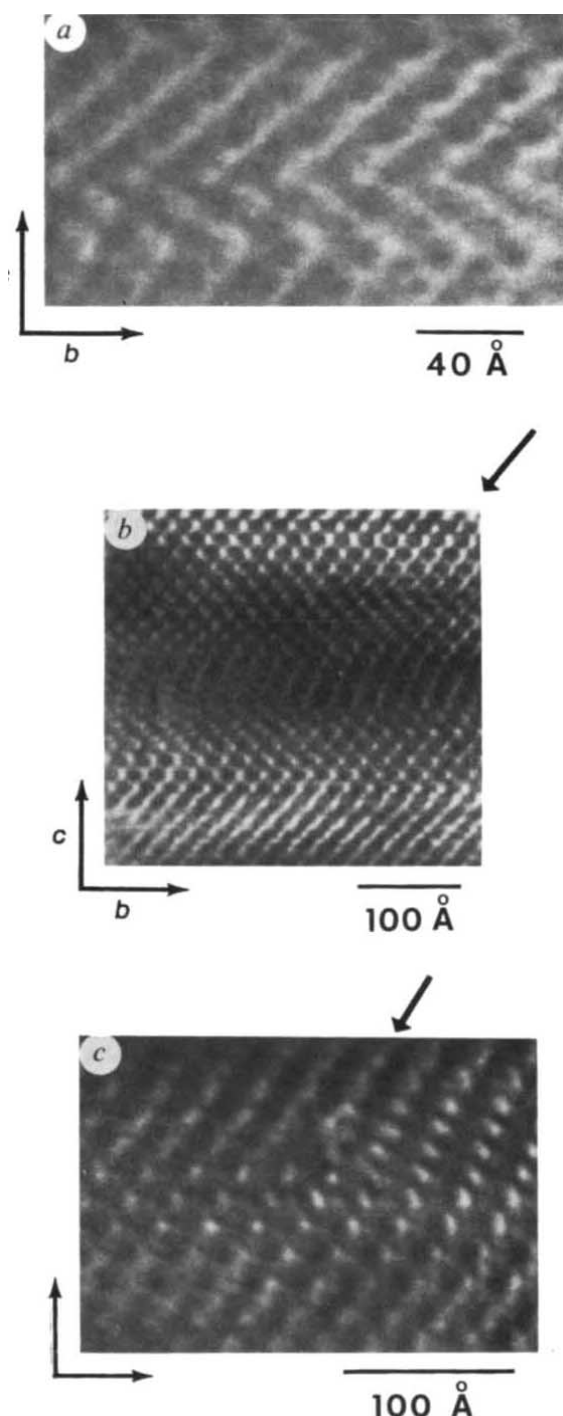


Fig. 5 Lattice images showing (011) fringes. Axes refer to ideal orthorhombic structure. The chevron structure in (a) is due to small-scale twinning of two monoclinic structures. Varying orientation of the fringes in (b), due to stacking disorder, and the dislocation in the superstructure in (c), with apparent projected Burgers vector $\frac{1}{2}[011]$, are best seen by viewing the images at a low angle in the direction of the arrows.

to date possesses pervasive structural disorder. The structure has occasional missing and extra sublayers. More important, we have shown that the average orthorhombic structure actually consists of a complex mixture of orthorhombic material with monoclinic material (probable space group $P2_1/n$) in two twinned orientations and having variable angles between the (001) and '010' fringes. The variable angles may result from variations in the stacking of rows of low metal occupancy or atomic number

in adjacent sublayers. It is also possible that the twinned monoclinic structure results from a phase transition from a higher-temperature orthorhombic structure.

We speculate that two different types of layers corresponding to the light (001) and '010' fringes may be responsible for the two components of superconductivity. The component with onset temperature ~ 120 K may result from local superconductivity within the '010' layers, while the lower-temperature component may result from superconductivity in the (001) layers. It is thus possible that this structure represents a three-dimensional superconductor.

This work was supported by NSF, NASA, the Texas Center for Superconductivity and the TLL Temple Foundation (Houston). We are indebted to Dr J. E. Post of the Smithsonian Institution for AEM analytical standards.

Received 11 February; accepted 3 March 1988.

1. Chu, C. W. *et al. Phys. Rev. Lett.* **60**, 941-943 (1988).
2. Hazen, R. M. *et al. Phys. Rev. Lett.* **60**, 1174-1177 (1988).
3. Michel, C. *et al. Z. Phys.* (submitted).
4. Aurivillius, B. *Arkiv Kemi* **1**, 463-480; 499-512 (1949); **2**, 519-527 (1950).
5. Livi, K. J. T. & Veblen, D. R. *Am. Miner.* **72**, 113-125 (1987).
6. Cliff, G. & Lorimer, G. W. *J. Microscopy* **103**, 203-207 (1975).

The mechanism of operation of tin(IV) oxide carbon monoxide sensors

P. G. Harrison* & M. J. Willett†

* Department of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK

† British Coal Corporation, HQ Technical Department, Ashby Road, Bretby, Staffordshire DE15 0QD, UK

The electrical properties of tin(IV) oxide render it an excellent material for the detection of very low levels of carbon monoxide, and several devices based on sintered pellets of SnO_2 are commercially available^{1,2}. Investigations of the operation of such sensors have generally been restricted to studies of the electrical behaviour of the oxide and little attention has been paid to the relationship between the surface adsorption phenomena and electrical changes in the bulk oxide. These studies³ demonstrated the consistency of a model in which conductance is effectively controlled by the population of negatively charged oxygen adsorbates. The chemical reactions occurring at the oxide surface, although unsubstantiated, have been assumed to involve CO adsorption, desorption of CO_2 (which produces an increase in conductance) and replenishment of the resulting surface oxygen vacancy by adsorption of molecular oxygen⁴. Here we describe experiments in which the nature of the surface species adsorbed into tin(IV) oxide from atmospheres containing low levels of carbon monoxide is monitored using transmission infrared spectroscopy while simultaneously measuring the electrical changes produced in the bulk oxide. The data demonstrate the effect of heat treatment on the formation of adsorbate species and the mutual interdependence of adsorbate species and bulk oxide electrical conductance. This relationship is important both for sensor operation and heterogeneous catalysis.

The tin(IV) oxide sample used for each experiment comprised a thin self-supporting pressed disk of high-surface-area gel (prepared by the hydrolysis of tin(IV) chloride). Two electrodes attached to the disk surface using gold paste allowed the electrical conductance to be continuously monitored. The sample was mounted in a twin-beam infrared cell provided with an integral furnace, which was positioned in the optical path of a dispersive infrared spectrometer and connected to a high vacuum system. Using this arrangement the oxide disk could be preheated at a temperature up to 660 K, exposed to atmospheres of any desired composition or pressure (usually 10^5 Pa), and the gas phase could be removed rapidly by evacuation. All

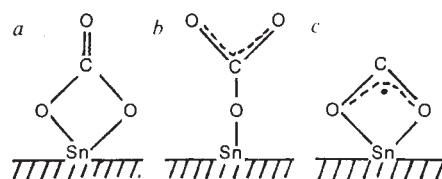


Fig. 1 Structures of the surface species formed on tin(IV) oxide from CO/air mixtures. *a*, Bidentate carbonate (characteristic infrared bands; 1,585, 1,223 cm^{-1}). Unidentate carbonate (1,430, 1,370 cm^{-1}). *c*, Carboxylate (1,540, 1,300 cm^{-1}).

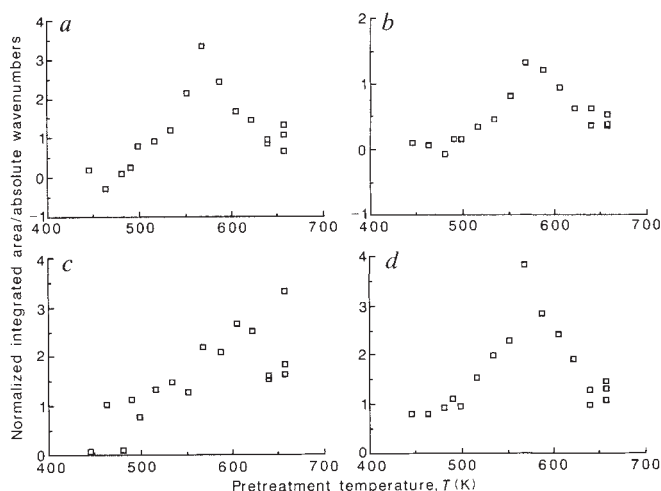


Fig. 2 Plots of the integrated peak areas against oxide pre-treatment temperature for the bands at 1,585 cm^{-1} (*a*), 1,223 cm^{-1} (*b*), 1,300 cm^{-1} (*c*), and 1,430 cm^{-1} (*d*).

the data reported here were obtained with the oxide sample held at 329 K.

Infrared difference spectra due to surface adsorbed species were obtained by subtracting the oxide absorptions from those observed in the oxide plus adsorbate system. The spectra observed on dosing the tin(IV) oxide disk with 1,000 p.p.m. CO in dry air comprise five major bands corresponding to three types of surface species: surface unidentate and bidentate carbonate, and a surface carboxylate (Fig. 1). The intensities of all the bands due to the adsorbate species increase steadily with time during the period of exposure. Removal of the CO/air atmosphere by evacuation followed by re-exposure to dry air produces no new infrared features but the bands due to surface unidentate and bidentate carbonate are greatly reduced in intensity. In contrast, those due to surface carboxylate are only slightly affected. It would appear that whereas both the surface carbonate species are readily desorbed, the surface carboxylate species are held more strongly.

The pre-treatment temperature of the oxide has a marked effect on the abundances of the three types of surface species formed. The absorption intensities (and hence the surface population) of both of the surface carbonate features increase with increasing calcination temperature up to ~ 570 K and decline rapidly thereafter. The absorption peaks for carboxylate, however, increase fairly steadily up to at least 620 K. These effects are clearly demonstrated by evaluating integrated peak areas for the characteristic peaks of each species, and in Fig. 2 these are plotted against oxide calcination temperature for the 1,223 cm^{-1} , 1,585 cm^{-1} (both due to bidentate carbonate), 1,430 cm^{-1} (unidentate carbonate), and 1,300 cm^{-1} (carboxylate) features respectively. Clearly the nature of the adsorbed species formed from CO/air atmospheres varies significantly depending upon the conditions, and at higher pre-treatment temperatures

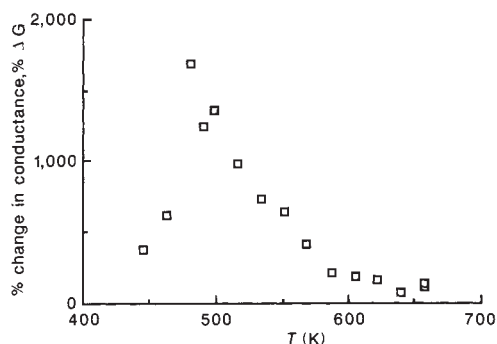


Fig. 3 Plot of the percentage conductance change $\% \Delta G$ against oxide pre-treatment temperature.

the main surface species formed by adsorption from CO/air atmospheres is the more strongly held surface carboxylate.

It might be expected that changes in the nature and population of the adsorbate species would affect the conductance behaviour of the oxide, and hence its efficacy as a sensing material. The change in absolute conductance ($\Delta G = G_{\text{CO/air}} - G_{\text{air}}$), however, exhibits essentially the same monotonic increase with oxide calcination temperature as the absolutely dry air conductance (G_{air}), rising to a peak at ~ 550 K. Thus the degree of thermal pre-treatment is the major factor controlling the absolute conductance levels and the effects attributable to the presence of CO are relatively minor. For this reason, the parameter normally used to express the performance of practical sensing devices is neither the absolute conductance (G) nor the conductance difference (ΔG), but rather the sensitivity or percentage conductance change ($\% \Delta G = 100 \Delta G / G$). When the data of the present study are expressed in terms of this parameter the dependence upon pre-treatment temperature exhibits a sharp peak at 490 K (Fig. 3), in contrast to the behaviour of ΔG . The magnitude of the change in $\% \Delta G$ is very large, the observed maximum corresponding to an increase in conductance of ~ 17 -fold compared to the value in dry air alone. The difference in the apparent sensitivity of the two parameters ΔG and $\% \Delta G$ to the presence of CO merely reflects the fact that ΔG is a measure of the average number of extra charge carriers produced by the presence of CO, whereas $\% \Delta G$ represents the significance of this number relative to the oxide electrical properties in the absence of CO.

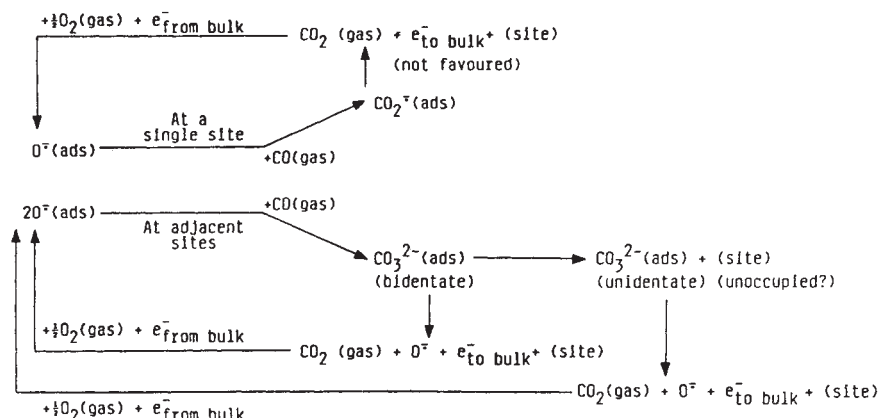
These combined infrared and conductance data may be described mechanistically by the scheme shown in Fig. 4. The most probable adsorption sites for CO under the conditions pertaining in this study are surface O_{ads}^- species⁵. The reaction of CO with adjacent pairs of O_{ads}^- produces surface bidentate carbonate which can subsequently be transformed into surface unidentate carbonate (plus a surface oxygen vacancy). Adsorp-

tion of CO at a single O_{ads}^- site will yield surface carboxylate. It should be noted that the formation of all three surface species occurs without any electron transfer to the bulk solid, and hence no conductance change would be expected if the reaction terminated at this stage. However, desorption of CO_2 as a product provides a mechanism by which electrons are returned to the conduction band of the oxide, from which an increase in conductance would be expected. The surface oxygen vacancy produced by this process is replenished by (dissociative) subsequent adsorption of molecular oxygen⁶. The infrared spectra demonstrate that the desorption of surface carboxylate (to CO_2 or any other product) is clearly not favoured, and also show that the abundances of the three types of adsorbate are highly dependent upon the nature of the oxide surface. The high hydroxylation of the initial oxide surface⁷ precludes a significant concentration of O_{ads}^- adsorption sites, and so all species are sparse. As the concentration of O_{ads}^- increases due to increasing dehydroxylation with increasing pre-treatment temperature, so the concentration of all three species increases, the adsorption at adjacent O_{ads}^- sites to give surface carbonate being favoured. At pre-treatment temperatures > 570 K, however, the specific surface area of the material begins to fall very rapidly⁷, adjacent pairs of O_{ads}^- sites become less frequent and surface carbonate formation is reduced whereas surface carboxylate formation is favoured. Thus, very little CO adsorption is expected at low pre-treatment temperatures because of high surface hydroxylation; at intermediate temperatures when dehydroxylation is advanced (but before rapid surface area loss) and the concentration of O_{ads}^- is high, adsorption to give surface carbonate is the predominant process; but at high pre-treatment temperatures, the abundance of O_{ads}^- sites is low, and adsorption to give surface carboxylate dominates.

The data obtained in this and other⁷ studies indicate that the nature of the SnO_2 surface is defined by its thermal history. When the oxide is exposed to air containing CO at 329 K, the equilibrium surface coverage of carbonate species is maximized on a surface pre-treated at ~ 570 K. Even though such groups are relatively easily lost from the surface by desorption of CO_2 (thus contributing to the conductance increase observed under CO), there is not a simple correlation between the carbonate coverage and the CO sensitivity as defined by $\% \Delta G$, which peaks at a pre-treatment temperature of ~ 490 K. This observation is readily rationalized by the fact that the infrared data represent the number of occupied O_{ads}^- sites where reaction with CO has occurred, whereas $\% \Delta G$ is a measure of unoccupied O_{ads}^- sites which are available for adsorption. The conditions favouring each of these processes will not necessarily coincide, and could, as observed here, give rise to a difference in the optimum pre-treatment temperature.

At 329 K, the desorption of surface carboxylate groups is not favoured, and the accumulation of such species offers one possible mechanism for inhibition of the sensing action of the oxide at these temperatures. Can a similar process be invoked for the

Fig. 4 Summary of favoured surface reactions on SnO_2 at 329 K under 10^5 Pa of dry O_2/N_2 , 1,000 p.p.m. CO.



case of a working sensor that operates at ~ 600 K? If the desorption of carboxylate is less facile than that of carbonate at these higher temperatures, then a similar poisoning mechanism may be postulated. Experimental observations analogous to those reported here, but made at higher temperatures, are desirable in order to test this hypothesis.

M.J.W. thanks British Coal for their support and for permission to publish this paper. The views expressed are those of the authors and not necessarily those of the British Coal Corporation.

Received 9 December 1987; accepted 2 February 1988.

1. Watson, J. & Yates, R. A. *Electronic Engng.* 47-57 (1985).
2. Karpel, S. *Tin and its Uses*, No. 149, 1-3 (1986).
3. McAleer, J. F., Moseley, P. T., Norris, J. O. W. & Williams, D. E. *JCS, Faraday I Trans.* 83, 1323-1346 (1987).
4. Windischmann, H. & Mark, P. J. *Electrochem. Soc.*, 126, 627-633 (1979).
5. Joly, J. P., Gonzalez-Cruz, L. & Arnaud, Y., *Bull. Soc. chim. Fr.* 11-17 (1986).
6. Chang, S.-C. *J. Vac. Sci. Technol.*, 17, 366-369 (1980).
7. Harrison, P. G. & Guest, A. *JCS, Faraday I Trans.* 83, 3388-3397 (1987).

Volatilization of tin as stannane in anoxic environments

Olivier F. X. Donard* & James H. Weber

Chemistry Department, Parsons Hall, University of New Hampshire, Durham, New Hampshire 03824, USA

Recent reports of volatilization of metals from the aquatic environment demonstrate a previously unsuspected loss from the aquatic metal budget. Volatilization can occur by various mechanisms. Arsenic volatilizes from the aquatic environment as trimethyl- and dimethylarsine formed by biomethylation¹; plants and soils form volatile selenium compounds²; volatilization of mercury from the oceans might rival that of anthropogenic origin³; maritime air can contain a significantly higher fraction of alkyl-lead compounds to total lead than continental or urban air⁴. Finally, increasing evidence suggests the potential importance of tin volatilization through formation of tetramethyltin by chemical or biological methylation in the environment⁵⁻⁷. We recently described possible formation of tetramethyltin by the macroalgae *Enteromorpha* sp. under oxic conditions⁶. Here we examine the contribution of decaying algal material to environmental cycling of tin by describing formation of volatile stannane (SnH_4) in anoxic environments.

We started with the assumption that decaying algal material contributes to methylation of tin in anaerobic, tin-amended systems through the release of known alkylating compounds. The ultimate result of several methylation steps would be the formation of volatile tetramethyltin. Methyl iodide (MeI) is a strong candidate for environmental methylation of tin. It is an abundant, naturally occurring metabolite, originating from seaweeds^{8,9}, which methylates Sn metal and Sn(II) in model reactions by oxidative addition processes¹⁰⁻¹². Results presented below indicate that under strictly anaerobic conditions, macroalgae induce volatilization of tin as stannane, not as tetramethyltin. This new reaction could be a pathway for loss of tin to the atmosphere in natural reducing environments such as salt marshes. The results suggest that a similar process also might be important for other metals and non-metals.

Six samples of 1.1 g of tissue-blotted *Enteromorpha* sp. macroalgae (Great Bay Estuary, New Hampshire) containing, depending on the specific sample, ~ 600 ng g⁻¹ inorganic plus organic tin⁶ were placed in 100 ml of artificial sea water in 198 ml reaction vials (Supelco, Inc.). There was no attempt to remove

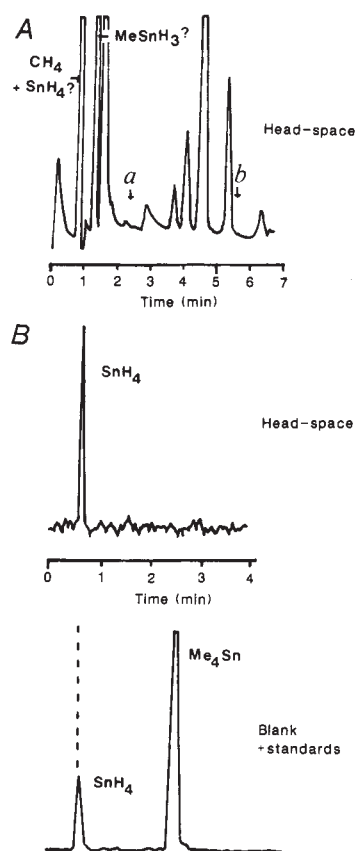


Fig. 1 Identification and quantization of tin compounds by GC and AAS in the head-space of sample 1 after six months. A, GC determination: parameters as given previously¹¹ except for a 40 ml min⁻¹ N₂ carrier gas flow. Stannane and methyltin standards have the following retention times: SnH₄, 0.64 min; MeSnH₃, 1.14 min; Me₂SnH₂, 2.13 min; Me₃SnH, 3.81 min; Me₄Sn, 5.59 min. Other retention times are: CH₄, 0.64 min; MeI, 2.46 min. Expected positions for MeI and Me₄Sn are denoted by a and b. B, AAS¹³ determination of SnH₄ and comparison with standards.

the epiphytic community. Vials 1 and 2 were unspiked. Vials 3 and 4 were spiked with 0.5 ml of 1,000 µg ml⁻¹ solution of SnCl₄ (Alfa, high purity). Vials 5 and 6 received 0.1 g of 99.9% tin metal (Alfa, 20 mesh powder). The vials were sealed with 'crimp-on' Teflon-lined silicone septa to allow measurement of the volatile tin compounds that might form. Solutions in the vials were purged for 20 min with argon to remove oxygen and stored at room temperature in the dark for six months. Monthly monitoring for the detection of any evolved volatile tin species was done by injecting 0.5 ml head-space subsamples into a gas chromatograph (GC)¹¹. Tin-specific identification of volatile tin species was done by injecting samples into our atomic absorption spectrometer (AAS) system^{13,14} through an injection port placed just before the chromatographic column.

Monthly gas chromatography (GC) measurements of compounds in head-space samples indicated the occasional occurrence of both MeI and CH₄ during the first three months. After six months, the head-space of all vials revealed a major peak whose retention time matched exactly that of MeSnH₃ standards. All vials presented a qualitatively similar gas distribution with greatest peak area for CH₄ and no MeI peak (Fig. 1A). The arrows show that no MeI (a) or Me₄Sn (b) occurred.

Determination of tin hydride compounds by the tin-specific AAS detector without addition of NaBH₄ (Table 1) demonstrates the presence of stannane in the gaseous and aqueous phases. Yields for tin-amended samples range from 0.4% for Sn(IV)-amended vials 3 and 4 to 10⁻³% for Sn⁰-amended vials

* Present address: Groupe d'Océanographie Physico-Chimique, Université de Bordeaux I, 33450 Talence Cedex, France.

Table 1 Concentration (ng ml⁻¹ as Sn) of stannane and methyltin compounds in gaseous and aqueous phases without addition of NaBH₄

Sample	Sample number	Head-space	Dissolved in the aqueous phase				
		SnH ₄	SnH ₄	MeSnH ₃	Me ₂ SnH ₂	Me ₃ SnH	Me ₄ Sn
Unspiked	1	1.07	4.4	n.d.	n.d.	n.d.	n.d.
Unspiked	2	1.33	4.9	n.d.	n.d.	n.d.	n.d.
Sn(IV)	3	2.37	20.6	traces	n.d.	n.d.	n.d.
Sn(IV)	4	1.81	18.1	n.d.	n.d.	n.d.	n.d.
Sn ⁰	5	1.55	21.1	n.d.	n.d.	n.d.	n.d.
Sn ⁰	6	1.46	21.2	traces	n.d.	n.d.	n.d.

Tin added: 500 µg SnCl₄ to samples 3 and 4; 0.1 g Sn⁰ to samples 5 and 6. Quantization done by AAS¹³. Head space samples (20 ml) were injected just before chromatographic trap by an anoxic syringe technique. Aqueous samples (20 ml) were purged in the reaction vessel and gaseous tin species were stripped by purging with He for 8 min; n.d., not detected.

5 and 6. Monomethyltin hydride, the only methyltin hydride observed, occurred only in trace amounts in the aqueous phase. The retention time of the only peak in the head-spaces exactly matched that of a stannane, but not a tetramethyltin, standard under similar conditions (Fig. 1B). The 20 ng and greater stannane amounts measured in 20 ml samples are well above the 0.9 ng limit of detection for the method¹³. The AAS experiment did not detect Me₄Sn (limit of detection = 30 pg, ref. 13). We are confident that the tin compound identified by AAS is SnH₄ because AAS is a highly specific tin detector^{13,14} and because no other compound of tin has a boiling point near -52 °C. Because the non-polar packing in the cold trap separates primarily by boiling point, no other tin-containing compound will have a retention time near 0.4 min found for stannane. MeSnH₃, for example, has a retention time of 1.1 min under the same operating conditions.

Addition of NaBH₄ to aqueous solutions confirms the presence of MeSn³⁺ and the occasional occurrence of di- and trimethyltin in tin-amended solutions (Table 2). Higher values

Table 2 Concentrations (ng ml⁻¹ as Sn) of total recoverable inorganic tin (TRISn) and methyltin compounds from the aqueous phase after addition of NaBH₄

Sample	Sample number	TRISn	MeSn ³⁺	Me ₂ Sn ²⁺	Me ₃ Sn ⁺	Me ₄ Sn
Unspiked	1	7.2	0.2	0.2	0.1	n.d.
Unspiked	2	7.5	n.d.	n.d.	n.d.	n.d.
Sn(IV)	3	38.5	6.8	1.5	0.9	n.d.
Sn(IV)	4	106.2	26.6	n.d.	6.6	n.d.
Sn ⁰	5	75.4	2.0	n.d.	1.7	n.d.
Sn ⁰	6	64.1	2.9	n.d.	17.5	n.d.

Tin added: 500 µg SnCl₄ to samples 3 and 4; 0.1 g Sn⁰ to samples 5 and 6. Quantization done by AAS¹³.

in Sn(IV)-amended solutions than in Sn⁰-amended solutions suggest a relatively higher efficiency for conversion of tetravalent tin rather than tin metal into monomethyltin. Monomethyltin does not primarily originate from the macroalgae alone because concentrations are so low in the absence of added inorganic tin (vials 1 and 2). These easily measured MeSnH₃ amounts contrast with trace amounts in the absence of NaBH₄ (Table 1). The ~7 ng ml⁻¹ total recoverable inorganic tin (TRISn) values in unspiked vials 1 and 2 are primarily contributions from ~600 ng g⁻¹ total tin in the macroalgae⁶. TRISn values for Sn(IV)-amended samples represent the relatively small fraction of added tin not associated with macroalgae.

Formation of volatile tin species under mild conditions in model systems strongly suggests their formation under environmental reducing conditions and importance in the biogeochemical cycling of tin in the environment. Jackson and

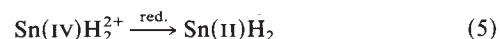
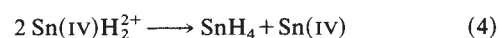
co-workers¹⁵ detected volatile methyltin and butyltin hydrides in Chesapeake Bay aqueous samples, but we know of no publication mentioning evidence for stannane. A major reason could be the low concentration in the aqueous phase due to its low (-52 °C) boiling point and the need for sensitive means of quantitation^{13,14}.

That carbon and tin are in the same group of the periodic table suggests that methanogenic bacteria can incorporate small amounts of tin while assimilating carbon and produce SnH₄. In fact, algae show analogous behaviour in methylating toxic arsenic taken up for phosphorus¹⁶. Methanogens, which are ubiquitous in the anoxic sediments, have the fundamental unifying characteristic of oxidizing H₂ and reducing CO₂ to form CH₄ (ref. 17). The presence of CH₄ in the vials confirms the presence of methanogens. Slow formation of SnH₄ in our vials agrees with the required removal of most or all SO₄²⁻ in artificial sea water before significant CH₄ production begins¹⁸. Evidence for methanogenic activity in our vials suggests that methanogens might be intimately involved in stannane formation.

We do not now have sufficient evidence to identify possible chemical or biological (enzymatic transformations) processes for production of stannane and methyltin compounds. Chemical reactions might result from biologically produced chemicals. A reasonable choice would be oxidative addition reactions of H₂ in which Sn⁰ and/or Sn(II) would be oxidized by two electrons per added H₂. The standard electrode potential of ~0.4 V at pH 7 demonstrates that oxidation by H₂ of Sn⁰ to Sn(II) and ultimately to Sn(IV) is strongly thermodynamically favoured (equations (1) and (2)).

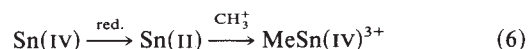


An uncertainty is that the Sn(II)H₂ intermediate is not a known compound. An alternative process beginning with Sn(II) (equation (3)) requires an intermediate rearrangement step (equation (4)) and/or reduction step as documented for arsenic¹⁶ (equation (5)) followed by reaction with H₂ (equation (2)).



We have insufficient data to distinguish between the above processes.

Relative solution concentrations of monomethyltin and trimethyltin depend on the form of tin added (Table 2). Monomethyltin predominates in Sn(IV)-amended solutions and trimethyltin predominates in Sn⁰-amended ones. Monomethyltin probably forms by oxidative addition of Sn(II) by a carbocation donor after its reduction from Sn(IV) (6).



Such a process occurs for monomethyltin formation in an Sn(IV)-amended anaerobic sediment¹⁹ and solutions containing macroalgae⁶. The methylating agent is probably MeI or 3-(dimethylsulphonio)propionate, both originate from algae⁶. Model studies^{11,12} confirm formation of monomethyltin from reactions of MeI and Sn(II). The observation that model reactions between MeI and Sn⁰ yield trimethyltin as a major product¹⁰ also agrees with trimethyltin formation in Sn⁰-amended samples (Table 2).

The prominent presence of SnH₄ in the gaseous and aquatic phases of the vials and the absence of Me₄Sn from these reactions in closed systems (Table 1) proves that stannane competes favourably with tetramethyltin in volatilization of tin in this model study. Particularly noteworthy are significant stannane yields in unspiked vials 1 and 2. This new volatilization pathway

could very well occur in reducing natural coastal environments such as salt marshes and microbial mats. It might compete with volatilization through Me_4Sn formation and could represent a previously unsuspected natural input of gaseous tin to the atmosphere. Because natural and anthropogenic sources of tin in the atmosphere are approximately equal⁷, stannane volatilization might be the origin of a significant fraction of total atmospheric tin. Relatively little is known about the stability of stannane, but our experiments¹¹ show that it is stable for at least several hours in the presence of water and oxygen. An important consideration is whether or not stannane can escape from the anoxic zone across the oxic layer into the atmosphere. This flux occurs for the analogous hydrogen-containing compounds methane²⁰ and easily oxidized hydrogen sulphide²¹ even though considerable oxidation occurs. Although we do not know the relative ease of oxidation of stannane compared to these two compounds, it is reasonable to assume some escape of stannane to the atmosphere.

One can conjecture about how important formation of hydride compounds is for volatilization of other metals and non-metals from the aquatic environment. The process is unlikely to be important for mercury or lead because of the thermal instability of their hydrides, or for arsine (AsH_3) because of its rapid reaction with air. In contrast, selenium, sulphur and germanium hydrides are reasonably stable and the macroalgal process dis-

cussed here for stannane might occur for these elements and others.

We thank Mark Hines for helpful comments. This research was partly supported by a National Science Foundation grant.

Received 30 September 1987; accepted 11 January 1988.

1. Mukai, H., Ambe, Y., Muku, T., Takeshita, K. & Fukuma, T. *Nature* **324**, 239–241 (1986).
2. Zieve, R. & Peterson, P. J. *Sci. tot. Envir.* **32**, 197–202 (1984).
3. Kim, J. P. & Fitzgerald, W. F. *Science* **231**, 1131–1133 (1986).
4. Hewitt, C. N. & Harrison, R. M. *Envir. Sci. Technol.* **21**, 260–266 (1987).
5. Craig, P. J. & Rapsomanikis, S. *Envir. Technol. Lett.* **5**, 407–416 (1984).
6. Donard, O. F. X., Short, F. T. & Weber, J. H. *Can. J. Fish. aquat. Sci.* **44**, 140–145 (1987).
7. Byrd, J. T. & Andreae, M. O. *Science* **218**, 565–569 (1982).
8. Lovelock, J. E. *Nature, Lond.* **256**, 193–194 (1975).
9. Gschwend, P. M., MacFarlane, J. K. & Newman, K. A. *Science* **227**, 1033–1035 (1985).
10. Craig, P. J. & Rapsomanikis, S. *Envir. Sci. Technol.* **19**, 726–730 (1985).
11. Rapsomanikis, S. & Weber, J. H. *Envir. Sci. Technol.* **19**, 352–356 (1985).
12. Ring, R. M. & Weber, J. H. *Sci. tot. Envir.* (in the press).
13. Donard, O. F. X., Rapsomanikis, S. & Weber, J. H. *Analyt. Chem.* **58**, 772–777 (1986).
14. Randall, L., Donard, O. F. X. & Weber, J. H. *Analyt. Chem.* **184**, 197–203 (1986).
15. Jackson, J. A., Blair, W. R., Brinckman, F. E. & Iverson, W. P. *Envir. Sci. Technol.* **16**, 110–119 (1982).
16. Andreae, M. O. in *Organometallic Compounds in the Environment* (ed. Craig, P. C.) 198–228 (Longman, London, 1986).
17. Zeikus, J. G. *Bacteriol. Rev.* **41**, 514–541 (1977).
18. Martens, C. S. & Berner, R. A. *Science* **185**, 1167–1169 (1974).
19. Gilmour, C. C., Tuttle, J. H. & Means, J. C. in *Marine and Estuarine Geochemistry* (ed. Sigleo, A. C. & Hattori, A.) 239–258 (Lewis, Chelsea, MI, 1985).
20. Bartlett, K. B., Harriss, R. C. & Sebach, D. L. *J. geophys. Res.* **90**, 5710–5720 (1985).
21. Cooper, D. J. et al. *Atmos. Envir.* **21**, 7–12 (1987).

The barite-opal-organic carbon association in oceanic particulate matter

James K. B. Bishop

Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York 10964, USA

Barite particles 0.5–5 μm in size are ubiquitous in the ocean and their formation, sinking and dissolution is a major part of the marine barium cycle^{1,2}. Barite formation appears to be caused by biological activity in the upper water column, but the exact mechanism is unknown. Analysis of 1–53 μm and >53 μm sized particles obtained by large volume *in situ* filtration in the Atlantic suggests that barites are formed in the >53 μm fraction in near-surface waters and released into the 1–53 μm fraction at depths below the euphotic zone. Scanning electron microscopy and energy dispersive X-ray fluorescence analysis of both size fractions shows first, that barites are formed almost exclusively in microenvironments containing decaying organic matter and the remains of siliceous plankton, and second that barites do not appear to be actively formed by the planktonic organism sampled. This explains the origin of suspended barite, and the similarity of dissolved Si and Ba distributions in the ocean. Suspended and sedimented barite may indicate the intensity of organic matter regeneration in the water column.

Barium has been extensively mapped in the world's oceans because of its surrogate role as a stable isotope of radium-226, a radioactive tracer of ocean circulation^{3–5}. A stable form is required as ²²⁶Ra is heavily involved in the biogeochemical cycle^{3,5,6}. In sediments, Ba occurs mainly as barite (BaSO_4) and is strongly enriched below equatorial zones of high biological productivity^{7,8}. A secondary source of sedimentary barite may be hydrothermal activity near mid-ocean ridges⁹. Recently, barite distributions in sediment cores have been proposed as an index of palaeoproductivity^{10,11}. Such applications require a better understanding of how Ba is involved in the biogeochemical cycle.

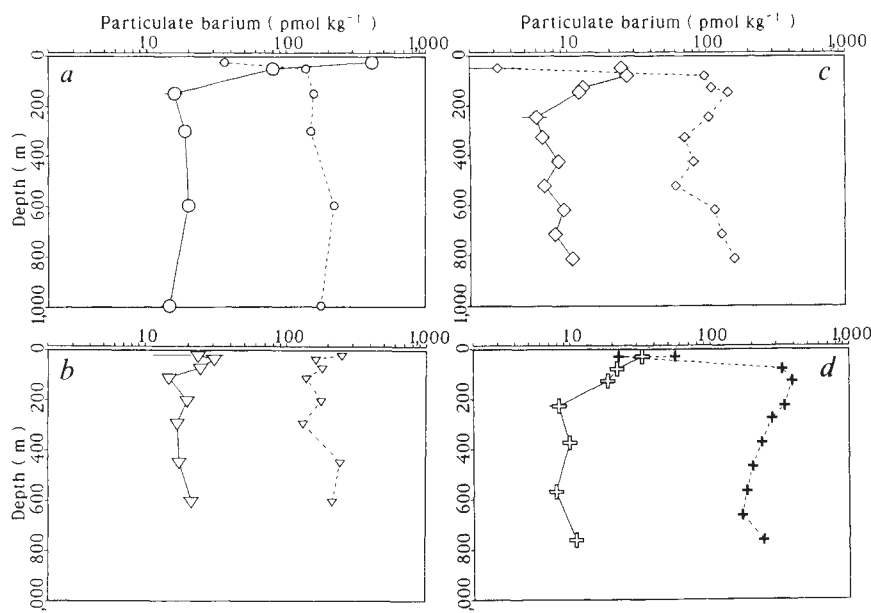
There has been much controversy about the mechanism by which biological activity causes barium to be transformed from

dissolved into particulate form. Hypotheses include: (1) formation of barite in microenvironments enriched with sulphate from decaying organic matter; (2) incorporation of barium into the skeletons of siliceous plankton; and (3) active precipitation of barite by other planktonic organisms. The first was based on seawater being undersaturated with respect to barite^{3,12}, and is supported by observations of barite in faecal debris¹. The second was suggested by the similarity of dissolved Ba profiles to those of the nutrient element Si⁶; by strong Ba enrichment in Black Sea plankton samples dominated by the diatom *Rhizosolenia*^{13,14}; and by the covariation of Ba and Si in estuaries¹⁵. This hypothesis was disfavoured when plankton samples from the Pacific Ocean¹⁶ were found to be 10–100 times lower in Ba than Black Sea plankton, and cultures of *Rhizosolenia* showed similarly low levels¹. A phytoplankton source was also ruled out as many particulate Ba profiles exhibit maxima at 100–200 metres, well below the euphotic zone^{1,2}. The third hypothesis^{1,17} was based mainly on the correlation between suspended barite concentration and primary productivity, and on the possible existence of planktonic protozoa related to the benthic protozoan *Xenophyophora*¹⁸, a known barite precipitator. Support also came from the discovery of the dominant role of the protozoan *Loxodes* in the freshwater barium cycle¹⁹. But a comparison of the freshwater *Loxodes* with its close marine relative *Remanella* showed that the marine form precipitated celestite and not barite²⁰. A planktonic barite precipitator has yet to be identified. This study seeks to clarify the mechanism of barite precipitation in seawater.

Particulate matter samples, divided into 1–53 and >53 μm size fractions, were collected by large volume *in situ* filtration²¹ from the upper 1,000 m of the northwestern Atlantic²². The >53 μm particle size fraction was not adequately sampled previously^{1,17} because samples were obtained by filtration of seawater collected by 30-litre Niskin bottles^{23,24}. Large volume filtration overcomes this bias as 1,000 times more water is filtered and samples are collected *in situ*. The 1–53 μm size fraction collected by large volume *in situ* filtration is comparable to the slowly sinking particle population sampled by the Niskin filtration method²³.

Most 1–53 μm Ba profiles showed the lowest concentrations in the euphotic zone (maximum depth, 70 m) and a large increase to high values by 100–200 m (Fig. 1a–d). Concentrations were lowest in waters of Sargasso Sea origin (Fig. 1c) and highest in

Fig. 1 Particulate barium distribution in 1–53 μm (small symbols) and >53 μm (large symbols) size fractions in the core waters of Warm-Core Ring (WCR) 82B on *a*, 18 June and *b*, 24 June 1982; *c*, newly formed WCR 82H, representative of the Sargasso Sea, 7 October 1982; and *d*, Slope Water, 15 October 1982. Although data from WCR is used in 3 of 4 examples, these profiles are a representative selection of those collected by large volume *in situ* filtration from the Slope Water, Gulf Stream and Sargasso Sea in the north-west Atlantic. Barium was extracted from filter subsamples in 0.6M HCl overnight at 60 °C and the resulting solutions were analysed by graphite furnace atomic absorption spectroscopy (J.K.B.B. manuscript in preparation). Blank corrections were minor. Relaching of leached samples demonstrated that barium extraction was complete. Error bars indicate the precision of each data point, which in most cases is smaller than the symbol.



the Slope Water (Fig. 1*d*), consistent with the biological productivity of these waters. This agrees with previous findings¹.

In contrast, >53 μm Ba was generally highest in the euphotic zone and decreased to uniform levels by 200 m. In some near surface samples, over 90% of the total Ba was in the >53 μm fraction. Replicate profiles taken one week apart (Fig. 1*a* and *b*) in the same water mass show that partitioning of Ba between 1–53 and >53 μm fractions can vary significantly in surface waters. All profiles (not just those shown) suggest that Ba is formed in >53 μm size fraction in shallow waters and released to the 1–53 μm fraction by 100–200 m. The Ba formation process appears to occur episodically.

Examination by scanning electron microscopy (SEM) and energy dispersive X-ray fluorescence (EDXRF) of several 1–53 μm samples from the 100–200 m Ba maximum showed mainly individual barite particles ranging from 0.5–5 μm in size. A small proportion of the barites contained minor amounts of strontium (Ba: Sr > 5), but most contained only Ba and S, again in agreement with Dehairs *et al.*¹. No barite was found in any intact organism.

The >53 μm sample from 24 m (Fig. 1*a*) was dominated by the diatom *Rhizosolenia* (uncountable because they occurred in dense mats in the sample) and the dinoflagellate *Ornithocercus* (850 per litre), and contained the highest amount of Ba. Barites occurred as clumps beneath organic and siliceous material (Fig. 2*a* and *b*) and also as clumps or strings of crystals inside broken *Rhizosolenia* (Fig. 2*c* and *f*). Individual members of such clusters were chemically and morphologically similar to individual barite particles found in the 1–53 μm fraction and to those described previously¹. Barites were not found in any structurally intact *Rhizosolenia*, or in association with broken or intact *Acantharia*, or with any of the numerous dinoflagellates in the same sample. Analysis of another Ba-rich >53 μm sample from 20 m in the Slope Water also showed the barite-opal-organic matter association.

The SEM/EDXRF study of barite-particle associations was extended to include some >53 μm samples collected from the southeastern Atlantic^{25,26}. The Ba results²⁵, although giving consistent profiles, showed little consistency with plankton group abundances or chemistry of the same samples. Sample selection was based on highest amount of Ba in the profile or on enrichment of specific organism groups. Like the North Atlantic profiles, the >53 μm Ba maxima were found in the euphotic zone. Samples were: C115-5-39N (R V *Chain* 115, Sta. 5, 40 m;

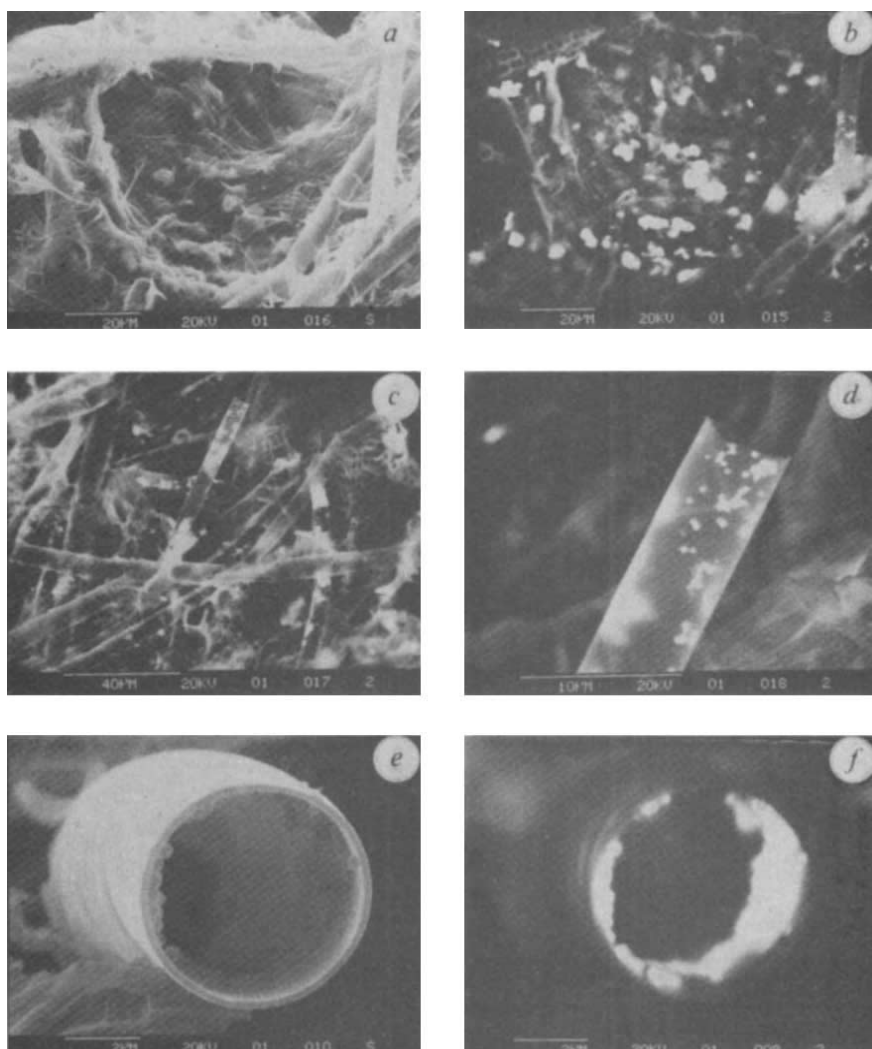
Ba = 900 pmol kg⁻¹) dominated by *Acantharia* and dinoflagellates; C115-5-44N (88 m; Ba = 80 pmol kg⁻¹), Foraminifera, broken and dissolving *Acantharia*; C115-6-47N (52 m; Ba = 123 pmol kg⁻¹), *Rhizosolenia*; C115-8-55N (20 m; Ba = 1,450 pmol kg⁻¹), centrate diatoms. No intact organism contained barite in any of these samples. Once again, barites dominantly occurred as clusters in association with organic matter and the remains of siliceous plankton.

The high concentration of barite in the *Rhizosolenia*-rich sample from the North Atlantic provides renewed credibility for the old Black Sea plankton data¹³. Also, plankton tows from the Pacific¹⁶ off Oregon described as “practically pure phytoplankton detritus” may have been misidentified *Rhizosolenia* mats containing comparable quantities of barite. The samples left a white precipitate (said to be SrSO₄) insoluble in concentrated HNO₃, HCl and HF (ref. 16) after sample extraction. Barite matches the chemical properties of the white precipitate better than SrSO₄ as the latter is easily soluble in the HNO₃ and HCl and barite is relatively insoluble. These samples were the most Ba-rich of all Pacific samples reported¹⁶ (320 p.p.m. Ba) and, when reanalysed later (J. Martin, personal communication), were twelve times more concentrated (3,800 p.p.m. Ba) than originally thought.

The occurrence of barite in broken *Rhizosolenia* and not in intact members of the same species reconciles the high Ba content in *Rhizosolenia*-rich plankton samples¹³ with low Ba content in actively growing cultures of this species¹. The consistent barite-opal association explains the observed first-order covariation of dissolved Ba and Si in the ocean and estuaries. The fact that secondary differences exist in profile systematics of these two elements^{2,6}, especially in the upper several hundred metres, suggests: (1) that some (but not all) siliceous plankton are important in the oceanic barium cycle and (2) that barite is formed in microenvironments rich in decaying organic matter and opal. This also explains why plankton group abundances and Ba are not well correlated in the water column.

These results support the hypothesis that barite precipitation in the water column occurs in microenvironments enriched in sulphate from decaying organic matter³. How much organically bound sulphur is required to react to form the observed barite? On a molar basis, diatom C:Si is about 2:1 and C:S is about 100:1, which yields a Si:S ratio of about 50:1. Because dissolved Si and Ba covary in a ratio of 1,300:1, the reaction of organically bound S to form barite in microenvironments needs to be only

Fig. 2 *a*, Secondary electron (SE) image $\times 1,000$ of barites beneath organic matter in the presence of opal. $>53 \mu\text{m}$ sample from 24 m (Fig. 1*a*). *b*, Backscattered electron (BSE) image of the same organic-silica rich environment. Barites here range from sub-micron to $5 \mu\text{m}$ in size. Barite particles appear bright against the organic carbon and silica of the sample as BSE intensity is determined by atomic number (z) of the sample. This greatly facilitated the search for Ba ($z = 56$) in samples dominated by elements C, O, Si and Ca (z ranging from 6–20). *c*, BSE image $\times 650$ showing the barites in degraded *Rhizosolenia* frustules but a lack of barite in intact members of the same species. *d*, BSE image. Transverse view of broken *Rhizosolenia* frustule at $\times 2,000$. *e*, Axial view (SE image) of same *Rhizosolenia* frustule $\times 10,000$. *f*, BSE image corresponding to *e*. Sample was carbon coated and analysed using a Cambridge SEM, beam energy 20 kV, with KEVEX energy dispersive X-ray fluorescence analysis system. Scale bars: *a*, *b*, $20 \mu\text{m}$; *c*, $40 \mu\text{m}$; *d*, $10 \mu\text{m}$; *e*, *f*, $2 \mu\text{m}$.



4% efficient if diatoms are the only sulphur source. As other plankton groups contribute to the particulate sulphur pool, the 4% estimate is an upper limit. The remains of siliceous organisms also appear to catalyse barite formation. Fresh silica surfaces may adsorb cations such as Ba better than other biogenic inorganic phases (J. Martin, personal communication), or diatom frustules may form better microenvironments than other plankton groups. Thus, it is concluded that conditions favouring barite formation (decaying organic matter and silica) are found in recently dead siliceous plankton and large aggregate particles, such as faecal material²⁶ and marine snow²⁷.

No support was found for the hypothesis that marine plankton (large or small) actively form barite. The use of barite as an indicator of organic matter regeneration intensity in the water column and sediments warrants further development. The diverse morphologies of suspended barite¹ could provide clues as to the mode of organic matter regeneration.

This research was supported by the National Science Foundation. M. D. Kurz, M. Q. Fleisher, D. Breger, R. F. Anderson, P. E. Biscaye and G. T. Shen are thanked for their contributions to the manuscript.

Received 7 October 1987; accepted 21 January 1988.

- Dehairs, F., Chesselet, R. & Jedwab, J. *Earth planet. Sci. Lett.* **49**, 528–550 (1980).
- Collier, R. W. & Edmond, J. M. *Prog. Oceanogr.* **13**, 113–199 (1984).
- Chow, T. J. & Goldberg, E. D. *Geochim. cosmochim. acta* **20**, 192–198 (1960).
- Wolgemuth, K. & Broecker, W. S. *Earth planet. Sci. Lett.* **8**, 372–378 (1970).
- Chan, L. H. *et al.* *Earth planet. Sci. Lett.* **32**, 258–267 (1976).
- Chah, L. H., Drummond, D. & Edmond, J. M. *Deep Sea Res.* **24**, 613–649 (1977).
- Revelle, R. R., Bramlette, M., Arrhenius, G. & Goldberg, E. D. *Geol. Soc. Am. Bull.* **62**, 221–236 (1955).
- Goldberg, E. D. & Arrhenius, G. O. S. *Geochim. cosmochim. acta* **13**, 153–212 (1958).
- Arrhenius, G. O. S. & Bonatti, E. *Progr. Oceanogr.* **3**, 7–22 (1965).
- Dymond, J. *Trans. Am. geophys. Un.* **66**, 1275 (1985).
- Schmitz, B. *Paleoceanogr.* **2**, 63–77 (1987).
- Church, T. M. & Wolgemuth, K. *Earth planet. Sci. Lett.* **15**, 35–44 (1972).
- Vinogradova, Z. A. & Koval'skiy, V. V. *Dokl. Akad. Nauk. SSSR* **147**, 217–219 (1962).
- Brongersma-Sanders, M. K. *Ned. Akad. Wet. Proc.* **B70**, 93–99 (1966).
- Li, Y. H. & Chan, L. H. *Earth planet. Sci. Lett.* **43**, 343–350 (1979).
- Martin, J. H. & Knauer, G. A. *Geochim. cosmochim. acta* **37**, 1639–1653 (1973).
- Dehairs, F., Lambert, C. E., Chesselet, R. & Risler, N. *Biogeochemistry* **4**, 119–139 (1987).
- Tendal, O. S. *Gaitea Rep.* **12**, 8–99 (1972).
- Findlay, B. J., Hetherington, N. B. & Davison, W. *Geochim. cosmochim. acta* **47**, 1325–1329 (1983).
- Rieder, N., Ott, H. A., Pfundstein, P. & Schoch, R. *J. Protozool.* **29**, 15–18 (1982).
- Bishop, J. K. B., Schupack, D., Sherrell, R. M. & Conte, M. H. in *Mapping Strategies in Chemical Oceanography* (ed. Zirino, A.) 155–176 (Am. Chem. Soc., Washington DC, 1985).
- Bishop, J. K. B. & Fleisher, M. Q. F. *Geochim. cosmochim. acta* **50**, 2807–2826 (1987).
- Bishop, J. K. B. & Edmond, J. M. *J. mar. Res.* **34**, 181–198 (1976).
- Gardner, W. D. *Limnol. Oceanogr.* **22**, 764–768 (1977).
- Kurz, M. D., Edmond, J. M. & Bishop, J. K. B. *Trans. Am. geophys. Un.* **58**, 1171 (1977).
- Bishop, J. K. B., Ketten, D. R. & Edmond, J. M. *Deep Sea Res.* **25**, 1121–1161 (1978).
- Allredge, A. L. & Youngbluth, M. J. *Deep Sea Res.* **32**, 1445–1456 (1985).

Irish tree rings, Santorini and volcanic dust veils

M. G. L. Baillie & M. A. R. Munro

Palaeoecology Centre, Queen's University, Belfast BT7 1NN, UK

There has recently been renewed interest in the dating of the violent eruption of the Aegean island of Santorini in the second millennium BC, both by its possible effects on tree-ring growth in the United States¹ (suggesting a date of 1628–1626 BC), and by acidity peaks in ice cores from South Greenland² (suggesting 1645 BC). We now show that oak trees growing on bogs in Northern Ireland produce significant concentrations of extremely narrow rings within a few periods less than 20 years long and that these periods correspond to the dates suggested by other methods for major volcanic eruptions. In particular, one of them, corresponding to a short period beginning in 1628 BC, was probably caused by Santorini. This date is qualitatively better than those derived from carbon-14 or ice cores, because it is based on an absolute tree-ring chronology.

In 1984 LaMarche and Hirschboeck¹ claimed that bristlecone pines growing near the upper tree line at several sites in western United States had been damaged by severe frosts, apparently caused by stratospheric dust veils produced by major volcanic eruptions. In particular they suggested that a significant frost ring, which was found in 1626 BC, might relate to the Santorini eruption and if so would date the eruption to that year or to 1–2 yr previously. More recently Hammer *et al.* have revised their original dating of Santorini³ of 1390 ± 50 BC, based on acidity peaks in Greenland ice cores, to 1645 ± 7 BC, based on new evidence from South Greenland². Both of these dates are in broad agreement with ¹⁴C evidence for a seventeenth century BC date for the eruption³.

The completion of the Belfast 7,272-yr oak-tree-ring chronology by 1984^{4,5}, and the subsequent archiving of the original ring-width measurements used in this chronology allow one to examine these data at specific dates. An initial survey of the 22 ring patterns that grew in the decade of 1620 BC showed that at least some trees put on narrow bands of rings at that time. In particular, two trees from Garry Bog in the north of Co. Antrim put on unmeasurably narrow rings in this decade. Trees from other sites showed very narrow rings at the same time. One tree, Q5392 from Sentry Hill, Co. Antrim, showed a colour change beginning in 1628 BC and another, Q1276 from Derrylard, Co. Armagh, showed anomalously small earlywood vessels (SEVs) (which elsewhere have been linked to anomalous weather conditions)⁶ in the growth ring for 1625 BC. This provides circumstantial evidence for an event in Irish bog oaks (oak trees that originally grew rooted in the peat of raised bogs), at a date suggested by other workers.

Figure 1 shows the ring patterns of four trees from 1626 BC. After 1628 BC there is a period when the rings in these trees are generally narrower than at any time in the previous century; furthermore, in each tree, one of these rings is the narrowest ring in the entire lifetime of the tree.

This coincidence in itself is a remarkable confirmation that the American frost-damage evidence represents a large-scale event, but is open to several objections. Some Northern Irish trees which grew at this time do not show such obviously narrow rings. There are other trees that have bands of narrow rings at dates when no major volcanic eruptions have been suggested. So one must survey the entire prehistoric period to see if oak growth was significantly impaired in the 1620s BC.

The narrow rings and narrow bands of rings close to 1626 BC suggested that some quantification of narrowness was needed. In theory the narrowest ring width (or widths where several rings are equally narrow) should be related to the worst growth conditions that the tree experienced during its lifetime. In Irish

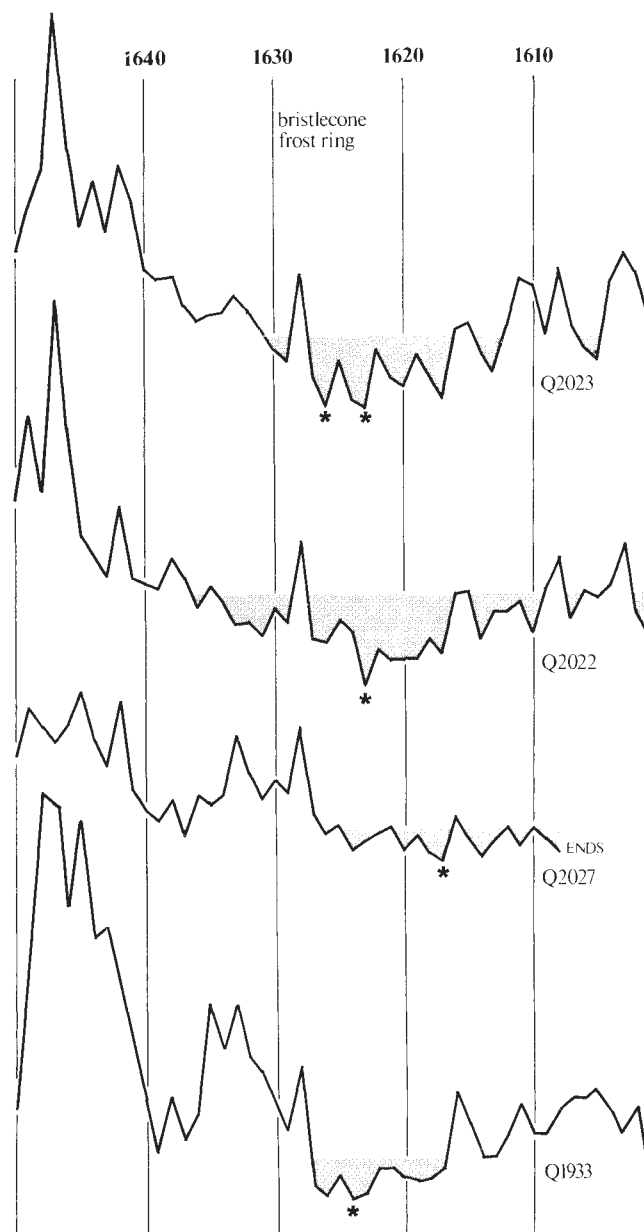


Fig. 1 Four trees from Garry Bog, Co. Antrim, Northern Ireland, showing curtailment of growth (the shaded areas are rings < 0.5 mm wide), associated with the narrowest rings in the lifetime of each tree (marked by asterisks). All the trees have more than 220 rings. The shaded square marks the American bristlecone pine frost ring dates for the eruption of Thera¹.

bog oaks the width of the band of spring vessels which comprise the early wood in each growth ring is approximately equal (and is conditioned by food reserves from the previous year). The overall ring width tends therefore to be a minimum in years where there is no summer growth and the width of the growth ring is merely the width of the spring vessels. But there is a relatively rare condition (SEVs)⁶ where the spring wood itself is reduced and disorganized; when this coincides with missing summer growth the tree may put on its narrowest ring. Because of the wide variety of possible causes of suppressed growth, the long lives of Irish oaks, and the fact that they were not all growing over the same period, one would expect the dates of the narrowest rings to cluster only to the extent to which cross dating between trees tends to align narrow rings.

Figure 1 shows that the dates of narrowest rings are spread over a period of as much as a decade within a particular narrow

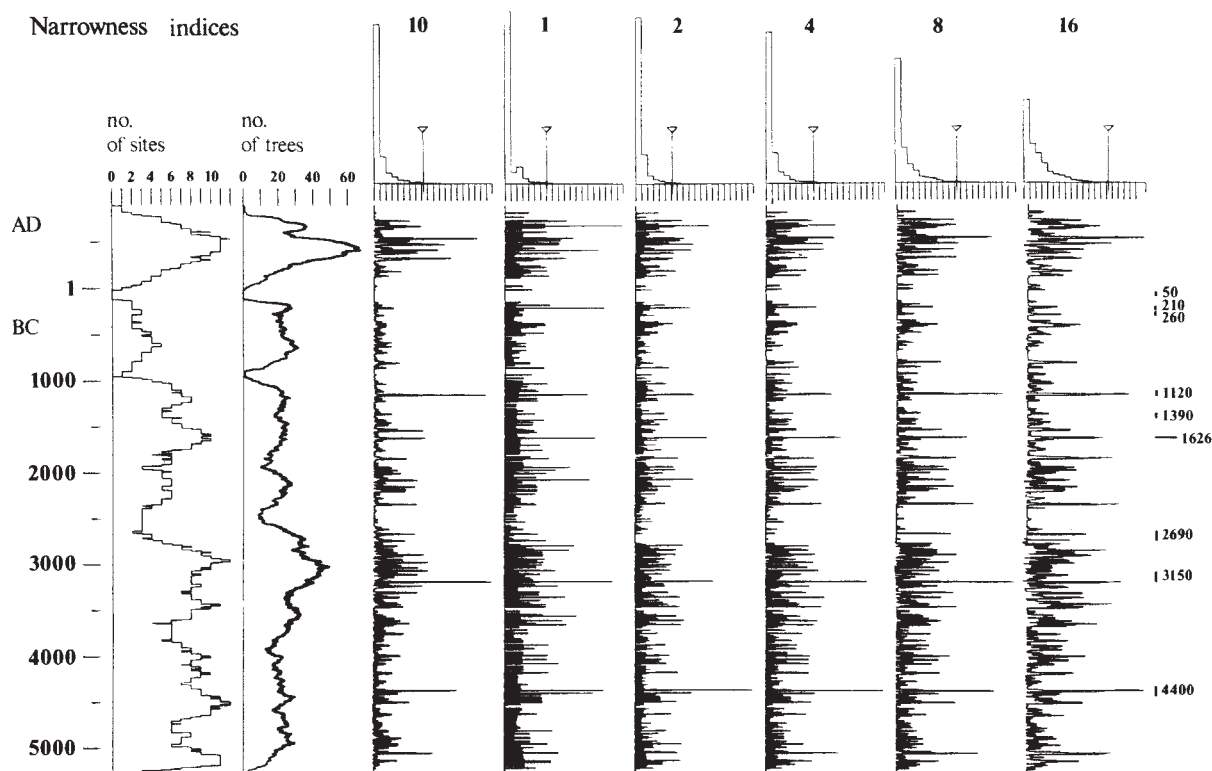


Fig. 2 The jagged vertical curves show the values of the $r \ln(r/\hat{r}) + s \ln(s/\hat{s})$ narrowness index for various window sizes, apart from the 10-yr window curve, which shows the simple rs index. The histogram above each curve shows the distribution of the index values, with the upper percentile marked by the arrow. The curves on the left give the total number of trees and sites in Northern Ireland used to produce the indices, and the dates on the right are of the acidity peaks originally identified in the ice core from Camp Century, Greenland³. More recent work^{2,8} gives a slightly different pattern of dates. The longer line shows LaMarche and Hirschboeck's frost ring date¹ in 1626 BC.

band, so we determined the total number of narrowest rings, r , within a period of fixed size passed over the chronology as a moving window. Because purely local effects, such as a landslide blocking the drainage of a bog, might cause all the trees at one site to put on a narrow band simultaneously, we determined the total number of sites, s , where at least one tree had its narrowest ring within the window period, and only considered patterns represented by at least two sites. Most of the trees in the earlier part of the chronology are from Northern Ireland, so for sake of consistency we only considered trees from this area. This means that although the chronology as a whole is continuous, there are two periods, 115–14 BC and AD 895–918, when there are no trees available from Northern Ireland². We paid most attention to the part of the chronology from 5289–116 BC, because this consists almost exclusively of bog oaks, whereas later parts contain increasing numbers of archaeological building timbers derived from a variety of environments; trees growing on peat must always have been vulnerable to adverse conditions, particularly flooding. We did not edit the archived data in any way, apart from correcting one case where multiple measurements of the same three trees had been mistakenly included, and restoring three trees that had been omitted precisely because they contained very narrow bands of rings.

As a preliminary test of this approach (using both the numbers of narrowest rings and of sites showing them) we ranked a very crude narrowness index, the product rs computed for a 10-yr window (see Fig. 2). We found that the three highest values in the prehistoric period, at 1153 BC, 3199 BC and 4377 BC, corresponded to three of the six major acidity peaks observed during this period in the Camp Century ice core³ (1100 ± 50 BC, 3250 ± 80 BC and 4400 ± 100 BC). Furthermore, the highest value in the period 13 BC–AD 894, at AD 541, corresponded to an acidity peak at $AD 540 \pm 10$. This demonstrated unquestionably that some eruptions were associated with clusters of narrowest rings,

and persuaded us to investigate better narrowness indices and different window sizes.

The rs narrowness index is proportional to the total number of sites n_s and rings n_r in the chronology at a particular date, so it can be improved by expressing the number of narrowest rings as a fraction of these, $rs/(n_r n_s)$, but we found that this exaggerated the importance of periods when there were few trees or sites in the chronology (because small random variations in absolute numbers produced large variations in the index at these periods); in addition, some trees put on more than one narrowest ring. A better index uses the expected number of narrowest rings and sites with narrowest rings rather than the total numbers within the window period.

The expected number of narrowest rings from a single tree in a particular window period is directly proportional to both the total number of narrowest rings in the tree, m , and the number of its rings that fall within the period, w , but is inversely proportional to the length of the tree, l . Thus a crude estimate of the narrowest rings from a single tree is just $w m / l$, and the estimated total number of narrowest rings, $\hat{r}$, is the sum of this quantity over all the trees. The probability that a tree will not put on a narrowest ring in a particular year is proportional to $((l-m)/l)$, thus an estimate of the probability that it will not do so during any year within the window period is given by $((l-m)/l)^w$, and the estimated probability that none of the trees from a site will put on a narrowest ring, $\hat{a}$, is the product of this quantity over all the trees from that site. The probability that a site does have at least one narrowest ring in the window period is $(1-\hat{a})$, so the expected number of sites, $\hat{s}$, is the sum of this quantity over all the sites.

The index $r \ln(r/\hat{r}) + s \ln(s/\hat{s})$ is based on the departures from these expected values, and includes weightings by the total number of trees and sites to diminish the effects of small sample sizes. Unfortunately one cannot compare it directly with a

theoretical distribution such as x^2 because s is dependent on r ; moreover because the trees in the chronology cross-date with each other they must be correlated, and some clustering of dates for narrowest rings should be expected, although the methods used to produce the expected numbers of trees and sites do not take this into account.

Figure 2 shows that major peaks in the index values occur at dates that previous workers^{2,5} have suggested for major prehistoric eruptions, from evidence that was not connected with tree rings. Such coincidences count as strong independent support for the eruptions dated at ~1150 BC, ~1626 BC, ~3195 BC and ~4375 BC. There is no supportive evidence for eruptions associated with our consistent index peaks at ~2345 BC and ~5060 BC, the former in particular appears to be a local single-site event. Interestingly our 207 BC single year peak suggests support for the 210 ± 30 BC ice-core evidence. Hammer *et al.*² have been unable to replicate the 1390 ± 50 BC peak seen at Camp Century (North Greenland) in the Dye 3 core (from South Greenland), and would now interpret it as a high latitude eruption whose acidic fallout did not extend to South Greenland. Although such an eruption could still have produced an extensive dust veil, one should not accept this until the acidity spike has been confirmed at other sites.

The narrowness index peak for the 1620s BC coincides with the 1628–1626 BC bracket of LaMarche and Hirschboeck. Inspection of the Northern Irish tree samples from this period suggests that they are variously affected from dates after 1630 BC. Although this differs slightly from the 1645 BC ice-core date, the tree rings must take precedence because every tree containing frost-damaged or narrow rings is absolutely dated. Such tree-ring dates cannot be shifted, even by a single year, without a complete redating of the chronologies that support them. This is impossible, because the Belfast chronology is strongly replicated internally and has been verified against other European chronologies⁵. The North American bristlecone pine chronologies are equally sound.

The ice-core date is expressed, however, as '... 1645 BC with an estimated standard deviation of ±7 yr, and an estimated error limit of ±20 yr'², so the tree-ring date already falls within the error limits. It also gives a result more in agreement with the ¹⁴C dates for the Santorini eruption discussed by Hammer *et al.*, falling well within the date range produced by calibrating the 2σ limits.

These results may have implications for interpreting the effects of volcanic dust veils. Even if volcanic ash clears from the stratosphere after 2–3 yr, the Belfast tree-rings show effects that last much longer (~10 yr), suggesting that an initial trigger event (flooding?) caused severe long-term problems for trees growing on bogs; other biological systems may show similarly extended responses. There are possible implications for the impact on human societies, which might suffer the effects of runs of bad harvests, poor pasturage and impeded communications. We draw attention to the 1150 BC event, which shows up as an extremely narrow band of rings in 90% of the trees covering the period. The effect begins dramatically in 1159 BC and recovery is not general until 1140 BC. 43% of trees on six sites have their narrowest rings during this period. There can be little doubt that this event is the Hekla 3 eruption observed by Hammer *et al.*³.

A perfectly plausible mechanism to explain how dust veils could cause adverse growth conditions on Irish bogs comes from Kelly and Sear's study⁷ of the climatic impact of explosive eruptions. They show that pressure changes in the northern hemisphere months after northern eruptions can produce a negative pressure anomaly over the British Isles. Such anomalies, associated with increased precipitation in middle latitudes, could provide exactly the conditions that might tip the balance against trees growing on marginal bog environments.

This work was begun by M.G.L.B. during a sabbatical term as Visiting Professor at the University of Arizona, Tucson.

Received 5 October 1987; accepted 4 January 1988.

1. LaMarche, V. C. Jr & Hirschboeck, K. K. *Nature* **307**, 121–126 (1984).
2. Hammer, C. U., Clausen, H. B., Friedrich, W. L. & Tauber, H. *Nature* **328**, 517–519 (1987).
3. Hammer, C. U., Clausen, H. B. & Dansgaard, W. *Nature* **288**, 230–235 (1980).
4. Pilcher, J. R., Baillie, M. G. L., Schmidt, B. & Becker, B. *Nature* **312**, 150–152 (1984).
5. Brown, D. M., Munro, M. A. R., Baillie, M. G. L. & Pilcher, J. R. *Radiocarbon* **28**, 279–283 (1986).
6. Tapper, M., Fletcher, J. & Walker, F. *Br. archaeol. Rep. (Int. Ser.)* **51**, 339–342 (1978).
7. Kelly, P. M. & Sear, C. B. in *Proc. 3rd Conf. on Climatic Variations and Symp. on Contemporary Climate: 1850–2100*, 178–179 (American Meteorological Society, 1985).
8. Hammer, C. U. *Jökull* **34**, 51–65 (1984).

The dependence of convection planform on mode of heating

Gregory Houseman*

Research School of Earth Sciences, Australian National University, GPO Box 4, Canberra ACT 2601, Australia

A fundamental property of a convecting fluid is its planform—the distribution in the horizontal plane of hot rising regions and cold sinking regions. For the Earth's mantle the planform might be visualized as a map of subduction zones, hotspots and possibly ocean ridges. Here I report numerical experiments of convection at high Rayleigh number which show a strong dependence of planform on heating mode. When heat generation is distributed uniformly through the box the preferred planform consists of an ensemble of time-dependent cold axial sinkers distributed in a hot diffuse upward flow. When half of the heat is generated within the box and the other half is input through the base, the preferred planform consists of an array of hot axial plumes and elongated cold sheets. In the former case the mean horizontal wavelength is about equal to the layer depth; for the latter it is about twice the layer depth.

Many previous two-dimensional convection calculations have been published^{1–7} in which the planform is necessarily prescribed as a series of alternating (often time-dependent) hot and cold sheet-type structures that extend into the third dimension. In general, such a planform is unstable at large Rayleigh numbers⁸ and would be replaced by a set of three-dimensional structures^{9,10}. The three-dimensional planform is generally also time-dependent; given structures may die out or coalesce with other structures, while new ones may form by instability of the upper or lower thermal boundary layers¹¹.

An approximate model of convection in the Earth's mantle is defined by the three equations that specify incompressible flow, conservation of momentum (infinite Prandtl number) and conservation of energy¹²:

$$\nabla \cdot \mathbf{v} = 0 \quad (1)$$

$$\eta \nabla^2 \mathbf{v} + \rho \mathbf{g} = \nabla p \quad (2)$$

$$\frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T = \kappa \nabla^2 T + \frac{H}{\rho C} \quad (3)$$

where $\mathbf{v}$ is the velocity field, ρ is the density, p is pressure, $\mathbf{g}$ is acceleration due to gravity, η is dynamic viscosity, T is temperature, κ is thermal diffusivity, H is heat generation per unit volume and C is heat capacity at constant pressure. The Boussinesq approximation is assumed; the physical properties are constant except that where the density appears in equation (2) it depends linearly on temperature:

$$\rho = \rho_0(1 - \alpha T) \quad (4)$$

where ρ_0 is the density at $T = 0$ and α is the coefficient of thermal expansion.

* Permanent address: Department of Earth Sciences, Monash University, Clayton, Victoria 3168, Australia.

Fig. 1 *a*, Schematic drawing of the $4 \times 4 \times 1$ box, with axes conventions used in the calculations. *b*, Horizontal cross-section ($z = 0.75$) flanked by two vertical sections ($x = 1, 3$), showing variation of temperature within the box after the initial convective instability from a random temperature perturbation. The vertical black lines show the intersection of the three planes, assembled as in *a*. For vertical sections, up is inwards. Colour indicates temperature: the average temperature of the box is red, grading to white for colder fluid and to black for hotter fluid.

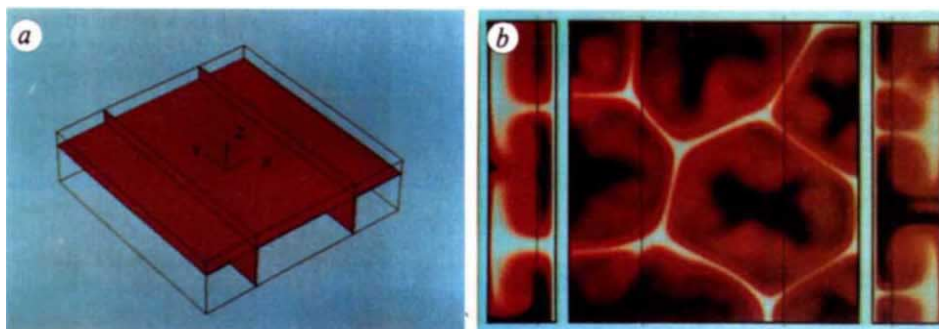


Fig. 2 Temperature sections (see Fig. 1 for conventions) at four different times, for the internally heated experiment ($\mu = 1$). Non-dimensional time is *a* 0.0911, *b* 0.1327, *c* 0.1630 and *d* 0.1945. Multiply by d^2/κ to obtain dimensional time. Note the absence of dark structures (hot jets or sheets).

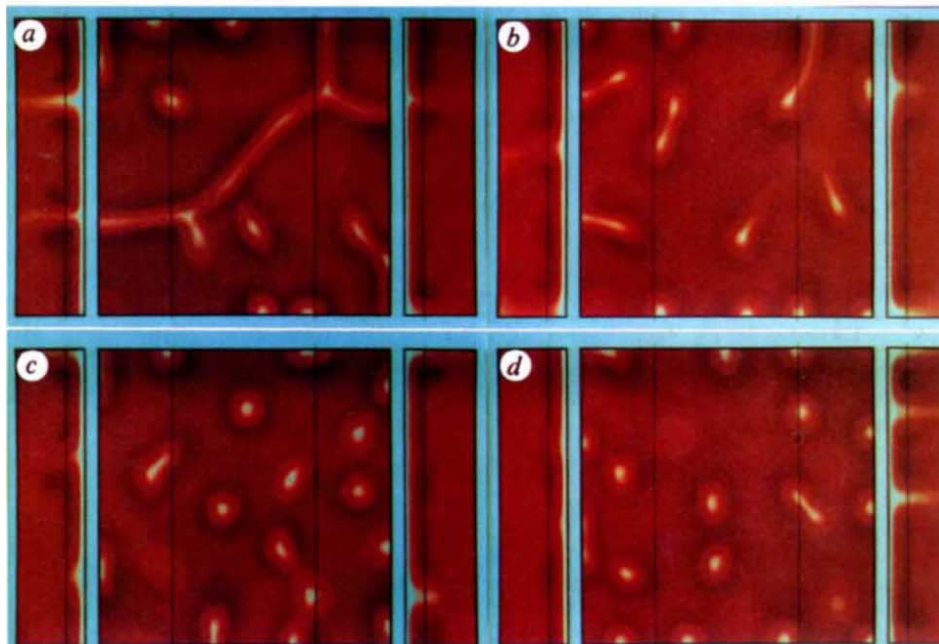
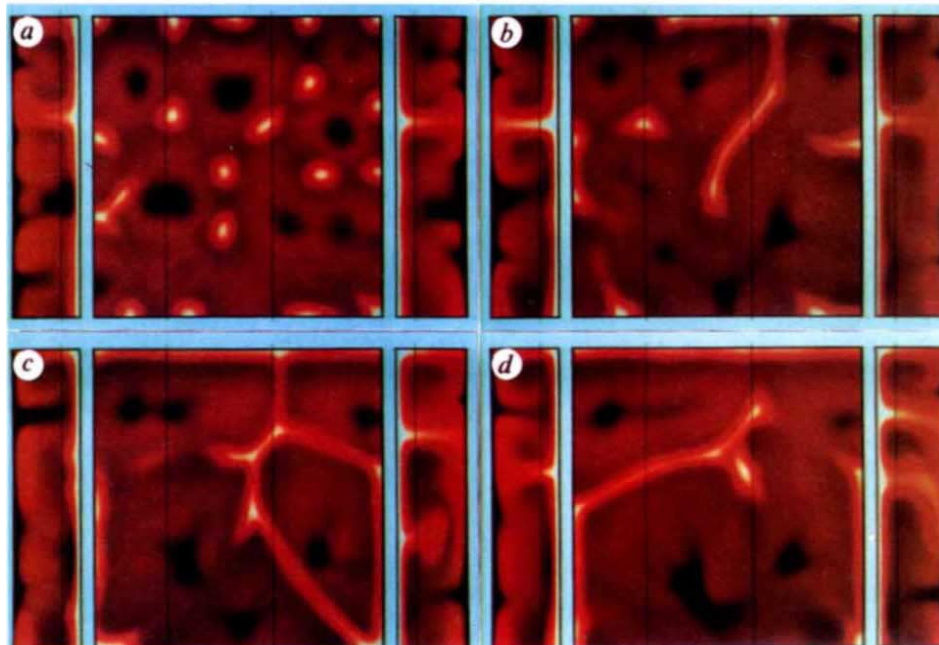


Fig. 3 Temperature sections at four different times for the experiment with half internal heating and half heating from below ($\mu = \frac{1}{2}$). Non-dimensional time is *a* 0.0117, *b* 0.0475, *c* 0.0898 and *d* 0.1202.



By Helmholtz' theorem, the velocity field can be represented as the curl of a solenoidal vector potential $\mathbf{A}$ (ref. 13):

$$\mathbf{v} = \nabla \times \mathbf{A} \quad (5)$$

and taking the curl of (2), one obtains

$$\eta \nabla^4 \mathbf{A} = \rho_0 g \alpha \left[\frac{\partial T}{\partial y} \hat{x} - \frac{\partial T}{\partial x} \hat{y} \right] \quad (6)$$

If there are no sources of vertical vorticity, the vertical component of $\mathbf{A}$ is zero, and for a given temperature field the above biharmonic equation is solved for each of the horizontal components A_x and A_y . The familiar two-dimensional formulation is obtained if either A_x or A_y is zero.

The equations are made dimensionless by a length scale d , a timescale d^2/κ and a temperature scale $(F + Hd)d/k$, where k is the thermal conductivity and F is the specified basal heat flow. The dimensionless equations are then completely specified by the Rayleigh number

$$R = \frac{\rho g \alpha (F + Hd) d^4}{\kappa k \eta} \quad (7)$$

and the heating mode number

$$\mu = \frac{Hd}{(F + Hd)}. \quad (8)$$

It is now possible, with the aid of a supercomputer, to solve the equations governing convection in three-dimensional planar geometry with relatively good resolution. The solutions presented here are for a constant viscosity fluid contained in a flat cell with stress-free upper and lower boundaries. The cell has a square cross section with an aspect ratio a (horizontal width: depth) of 4 (Fig. 1a). Reflection boundary conditions on the sides of the box require that the solution contains only horizontal wavelengths of the form $2ad/n$, where n is an integer. To permit a large range of wavelengths and to escape the artificial influence of the side boundaries, a should be large compared to 1, and in practice was as large as available computer memory permitted. The boundary conditions of zero normal velocity and zero shear-stress on the six surfaces of the box can be written as a set of conditions on the components (or their normal gradients) of streamfunction $\mathbf{A}$ and vorticity $\boldsymbol{\omega}$. Temperature is fixed ($T=0$) on the upper surface of the box ($z=1$) and a specified heat flux F is fixed² on the lower surface ($z=0$). The method of solution of the biharmonic equation has been described previously¹⁴. The temperature is advanced by timestepping, using (3) and the current velocity field.

For the numerical experiments described below $R = 587,235$ (1,000 times the critical Rayleigh number for the case $a=1$ and $\mu=0$). Although this value is probably smaller than that applicable to the Earth's mantle (estimated at between 10^6 and 10^9), it is about the largest value for which the calculations can be adequately resolved on the finite difference mesh used here ($129 \times 129 \times 33$ mesh points). Each of the experiments ran for 7,000 time-steps and required approximately 160 minutes of c.p.u. time on a Cyber 205 (two-pipeline, 32-bit precision). The experiments followed extensive testing and use of the program with $33 \times 33 \times 33$ mesh-point cubic boxes.

To initialize the first experiment (internally heated, $\mu=1$, Fig. 1b) the following temperature distribution was set for $t=0$:

$$T = T_1 \left[1 - z + 0.01 \sin(\pi z) \sum_k \sum_l \cos(kx + \phi_k) \cos(ly + \phi_l) \right], \quad (9)$$

consisting of a linear vertical temperature gradient perturbed by small harmonic terms that are randomly aligned in phase. The ϕ_k and ϕ_l are pseudo-random numbers between 0 and 2π , and the wavenumbers k and l range from $\pi/4d$ to $4\pi/d$, with increments of $\pi/4d$. The constant T_1 was estimated on the basis of previous two-dimensional calculations², so that the initial and final heat contents of the box are about equal, and the

evolution of the convection is not dominated by major gain or loss of heat from the box.

The 256 independent two-dimensional harmonics present in this temperature distribution grow at different rates in the initial convective instability. But in spite of the randomly chosen phase of each harmonic, the planform after a short time consists of approximately hexagonal cells that are surprisingly regular (see Fig. 1b). The sides of the hexagons consist of cold sinking sheets, and the centre of each cell consists of a broad, relatively hot, upwelling region. The distance between the centre of a cell and the side is about equal to the layer depth d . The degree to which the side boundaries influence the solutions is clear from the distortion required to force hexagonal cells into a square box. Note that the solution extends beyond the side boundaries by reflection in each boundary.

The hexagons are, however, a transient response to the initial highly supercritical temperature stratification of the layer. Breaks soon appear in the originally continuous network of cold sheets that form the sides of the hexagons and the broken sheets contract lengthways (Fig. 2a). With the contraction of some minor sheets, most of the downward flow takes place in a long elongated cold sheet extending across the box. The increased horizontal wavelength of the resulting cells results in another form of instability, and cold sinking blobs begin to form at a number of points in the upper boundary layer (Fig. 2a). The cold blobs develop into axial columnar structures that form a small scale circulation superimposed on a larger scale circulation in which the cold sheets nearly form a regular tessellation consisting of large octagons and small squares (Fig. 2a). The octagons are centred on two opposite corners of the box with the squares centred on the remaining two corners. Although such a tessellation has been observed in laboratory experiments¹⁵, the enforced symmetry of the side boundaries probably facilitates its formation here.

This planform also is unstable, and the elongated cold sheets that form the boundaries of the large scale cells continue to break and contract (Fig. 2b). As they do so the number of axial structures increases. Eventually, all of the extended sheets have contracted, and the planform consists entirely of cold axial columns separated by a relatively diffuse upward flow (Fig. 2c, d).

This final planform appears to be stable and is considered to be the preferred planform for an internally heated, constant viscosity, layer with zero heat-flux on the base and stress-free upper and lower boundaries. However, the flow remains strongly time-dependent and individual columnar structures appear to move horizontally with relative ease. The number of columns is maintained by an approximate balance between the rate at which they form (by instability of a thickening upper boundary layer, for example at $x=1, y=1$ in Fig. 2c) and the rate at which they disappear (by coalescing with their neighbours). The average surface density of cold columns varies between about $\frac{3}{4}$ and $\frac{7}{8}$ of a column per unit area (d^2), so the average distance between neighbours is typically slightly greater than d . Temporary elongated cold structures may form while two, or possibly three, cold columns are in the process of coalescing (see, for example, Fig. 2d, near $x=3, y=2$). But they generally contract back into axial structures relatively quickly. The planform observed in this experiment is similar to that obtained in a laboratory tank by Carrigan¹⁶, except that no large-scale component of flow is evident here.

In the second experiment ($\mu=0.5$, Fig. 3) the rate of internal heating is halved and an equal amount of heat is input by means of a fixed heat flow condition on the lower boundary. The final temperature distribution obtained from the previous experiment (Fig. 2d) was used as an initial condition for this experiment.

The convecting layer first responds to the change of heating mode by the formation of a thermal boundary layer on its base. This boundary layer is always much weaker than the upper thermal boundary layer (across which twice as much heat flows)

but it nevertheless generates hot rising plume structures (seen as dark areas in Fig. 3a) that are not present in the $\mu = 1$ case (Fig. 2). The array of cold axial structures present in the initial temperature distribution persists for a short time (Fig. 3a), although their horizontal mobility appears greatly reduced by the newly formed hot plumes. This planform, consisting entirely of axial structures (both hot and cold) is, however, unstable for this heating mode. The number of cold sinking structures is soon reduced as neighbouring axial sinkers join together to form elongated cold sheet structures (Fig. 3b). This behaviour differs from the coalescence of blobs characteristic of the $\mu = 1$ experiment, in that the resulting sheets appear relatively stable and do not contract back into axial structures.

Although the flow velocity is greater, the evolution of this planform is generally slower than in the previous experiment, probably because the cold sinkers are kept apart by the hot plumes present here. Cold axial structures continue to coexist with elongated sheets for some time (Fig. 3b) but the general trend is for the cold sheets to continue to increase in length by the joining up of neighbouring structures. In some cases these joins are unstable and the structures thus formed break up, only to be replaced by an alternative arrangement (Fig. 3c, d).

Throughout this experiment the hot plumes form a stable class of structures, and their distribution and number evolve in a manner analogous to that of the cold blobs of the preceding experiment. The average surface density of the hot plumes is, however, only about $\frac{1}{3}$ of a plume per unit area (d^2).

Towards the end of the experiment there are essentially four elongated cold sheets, three of which run along segments of the boundary ($x=0$, $x=4$ and $y=4$, Fig. 3d). The structure is occasionally modified by the formation of a cold blob that appends itself to one of the cold sheets. Although there might be further evolution of the planform if the experiment were continued further it would probably be limited to a rearrangement of the structures present at the end of the experiment. Relative to the $\mu = 1$ experiment there has been a significant increase in the horizontal wavelength of the convection planform (the average horizontal separation between neighbouring structures of the same type), and it is possible that the side walls of the container prevent attainment of an even longer wavelength.

In summary, the two experiments described above have demonstrated a strong dependence of convection planform on the mode of heat input, for a constant viscosity layer with stress-free upper and lower boundaries. The steady-state planform of the internally heated experiment ($\mu = 1$) is characterized almost entirely in terms of time-dependent cold columnar structures. With half of the heating from below ($\mu = 0.5$) the cold structures appear primarily as elongated sheets, with hot axial plumes also present. In both cases the planforms are time-dependent and have evolved through several stages, to a point where the nature of the planform has stabilized, although individual features continue to change. Under these circumstances, it is unlikely that the choice of initial conditions has determined the nature of the final planform¹⁷. The successive growth of different types of instabilities has eventually permitted selection of the most energetically favourable planform.

There is a clear need for further work, not only to consider the obvious cases of $\mu = 0$ and a larger container (for example $a=8$), but also for a systematic survey to show the type of planform expected for different types of boundary conditions (for example, stationary or rigidly moving, instead of stress-free, upper boundary) and the influence of properties such as temperature¹⁸ or depth-dependent¹⁹ viscosity, less-than-infinite Prandtl number, or spherical rather than Cartesian geometry.

With respect to convection in the Earth's mantle, the $\mu = \frac{1}{2}$ experiment is probably a closer approximation to reality than the $\mu = 1$ experiment; it is very likely that at least some of the heat input to the convecting layer comes from beneath²⁰. The resulting planform is also closer to what is inferred for the mantle circulation²¹. In a simple parallel, the cold sheet struc-

tures of the experiment are the subduction zones of the Earth and the hot plumes are the cause of the hot spot traces found on the Earth's surface. Hot sheet structures rising from the base of the layer beneath ocean ridges are neither required nor expected⁵. These parallels are perhaps coincidental in view of the many simplifications made in setting up the model (particularly the assumption of constant viscosity), but they suggest that the highly variable, possibly nonlinear²² viscosity of the Earth's mantle is perhaps not the major determinant of the mantle convection planform.

Finally, the timescale for the evolution of the planform in these experiments (the time unit $d^2/\kappa \approx 1.6 \times 10^{10}$ years even for $d = 700$ km) suggests that, for at least the first 10^9 years, the planform of convection in the mantle was similarly evolving away from an initial state that was the outcome of the processes of planetary accretion and core formation. The convection planform of the Archaean²³ may thus have had very little in common with the one evident at present.

Computing time on the CSIRO Cyber 205 was provided by a grant made under the CSIRO computing grants scheme.

Received 24 August; accepted 17 November 1987.

1. Richter, F. M. *J. Geophys. Res.* **78**, 8735-8745 (1973).
2. McKenzie, D. P., Roberts, J. M. & Weiss, N. O. *J. Fluid Mech.* **62**, 465-538 (1974).
3. Lux, R. A., Davies, G. F. & Thomas, J. H. *Geophys. J.R. astr. Soc.* **58**, 209-228 (1979).
4. Jarvis, G. T. & Peltier, W. R. *Geophys. J.R. Astr. Soc.* **68**, 385-424 (1982).
5. Houseman, G. A. *Earth planet. Sci. Lett.* **64**, 283-294 (1983).
6. Houseman, G. A. *Geophys. J.R. astr. Soc.* **75**, 309-334 (1983).
7. Davies, G. F. *Geophys. J.R. astr. Soc.* **84**, 153-183 (1986).
8. Busse, F. H. & Whitehead, J. A. *J. Fluid Mech.* **47**, 305-320 (1971).
9. Baumgardner, J. R. *Eos* **64**, 309 (1983).
10. Craig, C. H. & McKenzie, D. P. *Earth planet. Sci. Lett.* **83**, 123-136 (1987).
11. Howard, L. N. in *Proc. 11th int. Congr. Appl. Mech.* (ed. Gortler, H.) 1109-1115 (Springer, Munich, 1966).
12. Richter, F. M. *Rev. Geophys. Space Phys.* **11**, 223-287 (1973).
13. Richardson, S. M. & Cornish, A. R. H. *J. Fluid Mech.* **82**, 309-319 (1977).
14. Houseman, G. A. *Comp. Phys. Comm.* **43**, 257-267 (1987).
15. White, D. B. thesis, Univ. Cambridge (1981).
16. Carrigan, C. R. *Science* **215**, 965-967 (1982).
17. Foster, T. D. *J. Fluid Mech.* **37**, 81-94 (1969).
18. Richter, F. M., Nataf, H. C. & Daly, S. F. *J. Fluid Mech.* **129**, 173-192 (1983).
19. Gurnis, M. & Davies, G. F. *Geophys. J.R. astr. Soc.* **85**, 523-541 (1986).
20. Gubbins, D. *J. Geophys.* **43**, 453-464 (1977).
21. McKenzie, D., Watts, A., Parsons, B. & Roufousse, M. *Nature* **288**, 442-446 (1980).
22. Nakada, M. *J. Phys. Earth* **31**, 349-386 (1983).
23. Jarvis, G. T. & Campbell, I. H. *Geophys. Res. Lett.* **10**, 1133-1136 (1983).

Measurement of reduced peridotite-C-O-H solidus and implications for redox melting of the mantle

Wayne R. Taylor & David H. Green

Geology Department, University of Tasmania, GPO Box 252C, Hobart 7001, Australia

Methane-bearing C-O-H fluids may be important carriers of carbon and hydrogen in the Earth's deeper mantle^{1,2}. It has been proposed that such fluids have a key function in the genesis of kimberlitic magmas and diamond in the subcontinental mantle²⁻⁶. Because the melting behaviour of mantle peridotite in the presence of C-O-H fluids has not been accurately located as a function of pressure, temperature and oxygen fugacity^{1,4}, it has not been possible to assess the validity of reduced fluid 'infiltration' models of this kind. We report here advances in high-pressure fluid buffering techniques that have allowed the C-O-H fluid-saturated solidus of a fertile peridotite to be determined to 35 kbar. The results show that initial melting of peridotite saturated with H₂O-CH₄ dominated fluids occurs at temperatures well above stable continental geotherms, even under quite H₂O-rich conditions. This implies that C-O-H fluid-induced melting beneath continents will be restricted to regions where CH₄ can be largely eliminated from

the fluid by oxidation (redox melting) or where the geotherm is abnormally high (hot-spot melting), or where there is some combination of these processes.

The construction of realistic models for magma genesis in the upper mantle requires constraints on the mantle's compositional, thermal and redox structure, combined with knowledge of the melting relations of upper mantle materials (peridotite). Geophysical studies constrain the physical properties of the upper mantle and geochemical studies provide constraints on the mantle's chemical and mineralogical composition, but these two aspects can only be linked by direct experiment at high pressure and temperature. The current debate over the oxidation state of the Earth's mantle^{1,7-10} has led to the recognition that oxygen fugacity (fO_2) will be an important variable in magma genesis. Although the redox structure of the mantle is poorly known, a range of fO_2 s from above fayalite-magnetite-quartz to a lower limit near iron-wustite (IW) are probably appropriate^{1,5,9}. The nature of peridotite melting at the low fO_2 -end of this range, where CH_4 and H_2O are the dominant C-O-H volatiles and H_2 , C_2H_6 , CO_2 are minor components depending on the temperature (T), has been the subject of some speculation¹. In this paper we accurately locate the position of the peridotite-C-O-H solidus using a new high-pressure fluid buffering technique that allows simultaneous control of fO_2 and water activity (a_{H_2O}) on the graphite saturation surface at fO_2 near IW+1 log unit.

The fluid buffer consists of a mixture of α -WC, WO_2 and graphite (WCWO) generated by reaction *in situ* of W metal, WO_3 and C (mole ratio 29:1:60). C-O-H components are generated by pyrolysis of stearic acid ($C_{18}H_{36}O_2$). Peridotite of Hawaiian pyrolite composition minus 40% olivine¹¹ with a bulk Mg/Mg+Fe²⁺ ratio of 0.87 was prepared as a sintered oxide mix. Samples were loaded into inner graphite capsules of ~12 mg capacity and then sealed into Ag₅₀Pd₅₀ outer capsules of outer diameter 3.5 mm, together with ~170 mg of WCWO buffer mixture intimately mixed with sufficient stearic acid to generate ~3.5 mg C-O-H fluid. Talc outer sleeves were used in the high-pressure cell to maximize the external fH_2 environment. The internal capsule design is such that fluids freely permeate both the sample and buffer.

Experiments to define the peridotite-C-O-H solidus were performed in 0.5-inch piston-cylinder apparatus using techniques described previously¹¹. Run times varied between ~50 h for

experiments at 1,050 °C and ~10 h for 1,200 °C and higher. After each run the quenched fluid phase was analysed by a mass spectrometry/capsule-piercing technique, allowing CH_4/H_2O ratios to be determined. In all cases fluids were dominated by CH_4 and H_2O components. The capsule-piercing device and method of analysis has been previously described¹². Analyses obtained by this procedure are dependent on machine-related variables (principally ionization cross-section and secondary electron yield), making spectrometer calibration necessary to obtain quantitative CH_4/H_2O ratios. Because such ratios may also have been subject to modification during quenching from high T , calibration experiments were performed at 20 kbar, 1,100–1,200 °C, using furnace assemblies and techniques as close as possible to those used in experiments with peridotite. CH_4/H_2O fluids over the range 0.8–0.1 $CH_4/(CH_4+H_2O)$ were generated under pressure by thermal decomposition of *n*-hexacosane ($C_{26}H_{54}$) and $Al(OH)_3$ mixtures. Additional calibration experiments were performed at $fO_2 = IW$ (ref. 6). Results show that measured CH_4/H_2O ratios are too low by a factor of 2.0 ± 0.4 for $CH_4/(CH_4+H_2O)$ between 0.1 and 0.7. For $CH_4/(CH_4+H_2O) > 0.7$, the correction factor approaches unity.

In Fig. 1, fluid analyses from experiments with pyrolite at pressures from 9 to 35 kbar are plotted as a function of temperature and corrected $CH_4/(CH_4+H_2O)$ ratio. Water activities and reference buffer curves at 25 kbar were calculated from a modified Redlich-Kwong equation of state, calibrated for C-O-H fluid mixtures under upper mantle conditions of pressure and temperature². With the exception of four points that lie at very water-rich compositions, analysed fluids buffered by WCWO define a strongly temperature-dependent but pressure-insensitive line of best fit ($R = 0.94$) having the equation

$$\log fO_2 (\pm 0.2) = 11.82 - 33,780/T + 0.066(P/T)$$

where T is in Kelvin and P is pressure in bar. The equation incorporates a pressure term calculated from molar volume data. Relative to IW, WCWO lies ~0.5 log units above at 1,000 °C and ~1.3 log units above at 1,250 °C.

Experimental results for the system Hawaiian pyrolite-C-O-H are presented in Fig. 2. The subsolidus region is dominated by a large amphibole stability field that is truncated by the solidus at pressures less than 20 kbar. At $P < 20$ kbar amphibole persists to 15–20 °C above the solidus. Compositionally, amphiboles are

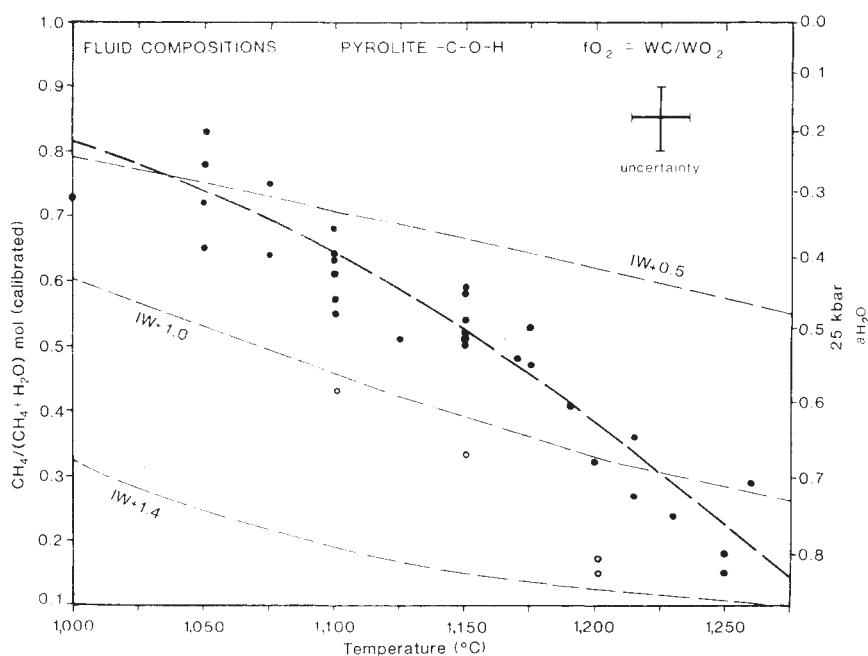


Fig. 1 Fluid compositions analysed by mass spectrometry plotted in terms of $CH_4/(CH_4+H_2O)$ ratio (calibrated) versus run temperature for pressures of 9–35 kbar. $CH_4/(CH_4+H_2O)$ ratios are independent of pressure within the limits of experimental uncertainty. Dashed line is the curve of best fit to the data points (●); open circles are data points excluded from the regression. Also shown are the 25-kbar reference buffer curves for IW+0.5, IW+1.0 and IW+1.5 log fO_2 units and calculated water activities for $P = 25$ kbar. The WCWO buffer is reproducible to within ± 0.2 log fO_2 units. There are no alloying problems that limit the usefulness of WCWO as encountered with buffers such as iron-wustite-graphite¹⁷. All fluid analyses were performed with a VG-micromass 7070F double-focusing mass spectrometer and out using a capsule-piercing device pre-baked at 150 °C and evacuated to 10^{-7} torr.

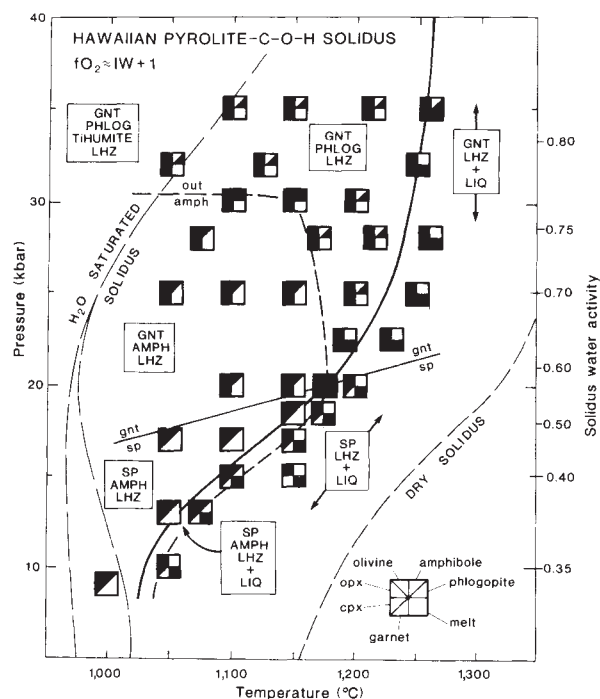


Fig. 2 Experimental results defining the solidus (heavy line) of peridotite of Hawaiian pyrolite composition in the presence of a C-O-H fluid buffered at $fO_2 \approx IW + 1$ log unit). All runs are fluid-saturated. Temperatures were recorded with Pt/Pt₉₀Rh₁₀ thermocouples controlled to within 5 °C of the set value. Size of the box symbols indicates the uncertainty in P and T . Heavy dashed line is the amphibole stability field. Light dashed line is the H₂O saturated solidus and accompanying amphibole stability field¹³. Calculated water activities along the solidus may be read from the right-hand axis.

pargasites with up to 30 mol% K-richterite component. Phlogopite is the dominant hydrous phase at $P > 30$ kbar and $T > 1,175$ °C, but is stable to no more than 5–10 °C above the solidus. Other minor subsolidus phases include titanoclinohumite and the oxides ilmenite and chromian rutile. Optical and electron microprobe identification of quenched melt and/or quench crystals and the absence of titanium oxides were used as the main criteria to define the solidus. Carbonate has not been identified in any run product.

The shape of the peridotite-C-O-H solidus differs from that found under water-undersaturated conditions where the solidus 'back-bends' near 30 kbar as a result of a large change in a_{H_2O} over a small pressure interval due to amphibole breakdown¹³. Because volatile activities are controlled and vary gradually and continuously, no back-bend of the peridotite-C-O-H solidus is observed. The solidi determined in both types of experiment are, however, related in that both follow paths in P - T - a_{H_2O} space. In buffered experiments this path may be quantitatively described if fluid compositions are known; in unbuffered experiments a_{H_2O} is undefined at pressures less than amphibole breakdown.

Using the regressed equation for WCWO and calculated activity coefficients for C-O-H fluid species², water activities may be determined along the experimental solidus (see Fig. 2). The graphite-present peridotite-C-O-H solidus surface may then be mapped out in P - T - a_{H_2O} space, given the bounds provided by the water-saturated solidus ($a_{H_2O} \sim 1$) (ref. 13) and volatile-absent solidus ($a_{H_2O} \sim 0$) of pyrolite (T. J. Falloon, unpublished results). The water-saturated solidus with graphite or diamond present (namely the water-maximum solidus) will lie at higher temperatures than those in ref. 13 due to dilution of the fluid

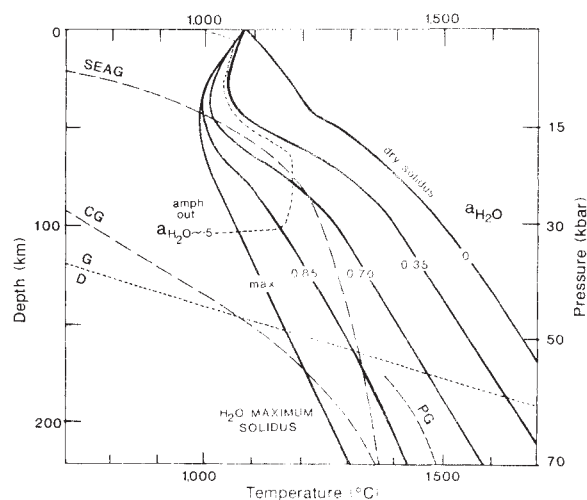


Fig. 3 Pressure-temperature projection of the P - T - a_{H_2O} solidus surface for the system peridotite-C-O-H (graphite/diamond present). Solidi for water activities of ~ 0 , 0.35, 0.70, 0.85 and the water maximum are shown. CG, continental shield geotherm¹⁸; PG, thermally perturbed geotherm¹⁸; SEAG, Southeastern Australian geotherm¹⁹. G, graphite; D, diamond. Light dotted curve is the amphibole stability field for $a_{H_2O} \sim 0.5$.

phase with both CH₄ and CO₂ (ref. 1). The decrease in a_{H_2O} is most pronounced at $P < 10$ kbar; at $P > 15$ kbar, a_{H_2O} exceeds 0.95 along the water maximum solidus.

Based on the above analysis, in Fig. 3 we present a projection of the C-O-H saturated solidus surface showing solidus curves for several water activities. We make the assumption that the solubility of solid components in the dominantly CH₄-H₂O fluids is not large enough to affect significantly the calculated activities. This is an acceptable approximation in the system peridotite-H₂O-CO₂ (ref. 14) and analogous behaviour is predicted for the reduced system.

At pressures less than 13 kbar the solidi in Figs 2 and 3 show large temperature depressions relative to the volatile-absent solidus, even though the fluid is CH₄-rich. Similar behaviour is observed in the system peridotite-H₂O-CO₂ at pressures where CO₂ is only weakly soluble in silicate melts^{1,15}. Between 13 and 20 kbar the peridotite-C-O-H solidi have an unusually shallow dP/dT slope that marks a change in the topology of the solidus surface between low and high pressures. It is in this region where fO_2 is controlled by WCWO that fluid compositions change rapidly from CH₄-dominated at low pressures to H₂O-dominated at high pressures (see Figs 1 and 2). Although fluids become H₂O-rich, their effect in depressing the peridotite solidus is not as marked as at lower pressures in that both CH₄ and H₂O appear to be important in determining melting temperatures. This behaviour could reflect a pressure-dependent increase in the solubility of reduced components in the melt phase^{12,16}.

At $P > 25$ kbar, the peridotite-C-O-H solidi show large temperature shifts away from the water-maximum solidus for only small dilutions of the fluid with CH₄. This shift is the opposite to that found at $P < 13$ kbar and is significant in that only solidi close to the water-maximum will be crossed by stable continental geotherms (CG in Fig. 3). For more CH₄-rich fluids ($a_{H_2O} < 0.9$), C-O-H solidi lie at temperatures well above CG. Thus these results clarify proposals for the generation of kimberlitic magmas by access of reduced C-O-H fluids from the deep mantle²⁻⁴. It is inappropriate to use peridotite-H₂O-CO₂ phase relationships in evaluating models calling on reduced C-O-H fluids because of the rapid changes in solidus temperature moving from oxidized, through a_{H_2O} tending to unity, to reduced CH₄,

H₂O fluids. Kimberlitic magmas will not be produced at low temperatures in the presence of CH₄-H₂O fluids. The caution that must be exercised in extrapolating from oxidized to reduced systems^{3,4} is therefore emphasized.

We may conclude from Fig. 3 that the involvement of reduced C-O-H fluids in kimberlite magmatism must be restricted to regions of the subcontinental mantle where there is significant thermal perturbation, such that temperatures lie on or above the kinked geotherm PG in Fig. 3. Alternatively, if a CH₄-bearing fluid phase undergoes sufficient oxidation, perhaps by interaction with high-*f*O₂ regions of the lithosphere^{2,16}, then aCH₄ will be lowered and aH₂O raised such that the solidus and geotherm intersect. We term these two processes 'hot-spot' style melting and 'redox' melting respectively.

Although not all petrologists favour open system fluid infiltration models to explain the origin of kimberlite, all would recognize the need for a source of volatile components. If this source is a fluid phase then reduced CH₄-bearing fluids will be the only suitable carriers of C and H in the deeper, subsolidus mantle^{1,2}. Our results demonstrate that quite water-rich C-O-H fluids can be stable in the subsolidus mantle at *f*O₂s below the water maximum (that is, $\approx 1W + 2 \log$ units, see ref. 2). The inflection in peridotite-C-O-H solidi at 80–100 km depth, however, will exclude reduced fluids from the shallow mantle by melting (for example intersection with high heat flow geotherms such as SEAG) and by amphibole crystallization (see Fig. 3). Thus evidence for the migration of reduced fluids is unlikely to be

preserved in spinel lherzolites as has been suggested by the results of recent *f*O₂ sensing studies⁷.

We acknowledge the technical assistance of N. Davies and K. Harris. This project was supported by a National Research Fellowship (W.R.T.) and the Australian Research Grants Scheme (D.H.G.).

Received 19 October 1987; accepted 27 January 1988.

1. Woermann, E. & Rosenhauer, M. *Fortschr. Miner.* **63**, 263–349 (1985).
2. Taylor, W. R. & Green, D. H. in *Kimberlites and Related Rocks* Vol. 1 (Geological Society of Australia, in the press).
3. Wyllie, P. J. *J. geophys. Res.* **85**, 6902–6910 (1980).
4. Wyllie, P. J. in *Kimberlites and Related Rocks* Vol. 1 (Geological Society of Australia, in the press).
5. Haggerty, S. E. *Nature* **320**, 34–37 (1986).
6. Foley, S. F. in *Kimberlites and Related Rocks* Vol. 1 (Geological Society of Australia, in the press).
7. Mattioli, G. & Wood, B. J. *Nature* **322**, 626–628 (1987).
8. Christie, D. M., Carmichael, I. S. & Langmuir, C. H. *Earth planet. Sci. Lett.* **79**, 397–411 (1986).
9. Arculus, R. J. A. *Rev. Earth planet. Sci.* **13**, 75–95 (1985).
10. Ulmer, G. C. *et al. Geol. Soc. Am. Spec. Pap.* **215**, 5–23 (1987).
11. Jaques, A. L. & Green, D. H. *Contr. Miner. Petrol.* **73**, 287–310 (1980).
12. Taylor, W. R. & Green, D. H. in *Magmatic Processes: Physicochemical Principles* (ed. Mysen, B. O.) 121–138 (The Geochemical Society, 1987).
13. Green, D. H. *Earth planet. Sci. Lett.* **19**, 37–53 (1973).
14. Schneider, M. E. & Eggler, D. H. *Geochim. cosmochim. acta* **50**, 711–724 (1986).
15. Mysen, B. O. & Boettcher, A. L. *J. Petrol.* **16**, 520–548 (1975).
16. Green, D. H., Falloon, T. J. & Taylor, W. R. in *Magmatic Processes: Physicochemical Principles* (ed. Mysen, B. O.) 139–154 (The Geochemical Society, 1987).
17. Esperancia, S. & Holloway, J. R. *Contr. Miner. Petrol.* **93**, 504–512 (1986).
18. Boyd, F. R. & Gurney, J. J. *Science* **232**, 472–477 (1986).
19. Griffin, W. L., Wass, S. Y. & Hollis, J. D. *J. Petrology* **25**, 53–87 (1984).

Extreme ¹⁸O depletion in calcite and chert clasts from the Elephant Moraine on the East Antarctic ice sheet

Gunter Faure*, Jochen Hoefs†, Lois M. Jones‡, John B. Curtis* & Douglas E. Pride*

* Department of Geology and Mineralogy and Byrd Polar Research Center, Columbus, Ohio 43210, USA

† Geochemical Institute, Goldschmidtstrasse 1, D-3400 Göttingen, FRG

‡ Conoco Incorporated, Ponca City, Oklahoma 74603, USA

Two large moraine es are presently forming near the margin of the East Antarctic ice sheet west of Victoria Land where the flow of the ice sheet is disturbed by subglacial bedrock ridges of the Transantarctic Mountains (Fig. 1). During the 1983–84 search for meteorites in this area several clasts, composed of black acicular calcite crystals, were collected from the Elephant Moraine (76°17'35" S and 157°20'05" E) where they are ablating out of the ice^{1–3}. Additional samples of this calcite were later collected from the Elephant Moraine during the 1984–85 and 1986–87 field seasons, but none were found at the Reckling Moraine at 76°15' S and 158°40' E^{4–6}. The calcite clasts are strongly depleted in ¹⁸O ($\delta^{18}\text{O} = -20.5 \pm 0.4\%$ with reference to standard mean ocean water, SMOW and in ¹³C ($\delta^{13}\text{C} = -22.6 \pm 0.1\%$, with reference to the PDB standard) in addition to being enriched in radiogenic ⁸⁷Sr ($^{87}\text{Sr}/^{86}\text{Sr} = 0.71417 \pm 0.00002$) relative to sea water. These results suggest that the calcite precipitated from aqueous solutions discharged by hot springs under the East Antarctic ice sheet.

The specimen shown in Fig. 2 (sample 582 in Table 1) was collected during the 1983–84 field season. It consists of black acicular calcite crystals aligned approximately at right angles to undulating layers. It resembles certain kinds of algal stromatolites that may be difficult to distinguish from calcrete^{1,7}. The black calcite also occurs as crusts on sedimentary material¹ and is less frequently associated with coarsely banded opaline chert ranging in colour from white to yellowish green⁴.

Table 1 Isotopic compositions of O, C and Sr in calcite and chert from the Elephant Moraine, Antarctica

Sample	Material	⁸⁷ Sr/ ⁸⁶ Sr	$\delta^{18}\text{O}$ (SMOW) (%)	$\delta^{13}\text{C}$ (PDB) (%)
582	Calcite	0.71417 ± 2	-21.1* -21.1* -19.75† -19.87† Average: -20.5 ± 0.4	-22.9* -22.8* -22.40† -22.47† -22.6 ± 0.1
583	Calcite	0.71402 ± 4	-14.2* -14.2*	-23.1* -23.0*
584	Calcite	0.71383 ± 3	ND	ND
EMA-2	Calcite	0.71488 ± 3	-12.91‡	-22.06‡
EMA-2	Chert	ND	+13.03‡	ND
F-84-6A	Calcite	0.71484 ± 1	-13.7*	-22.0*
Eimer and Amend SrCO ₃ standard		0.70809 ± 1	ND	ND

All ⁸⁷Sr/⁸⁶Sr ratios were corrected for isotope fractionation to ⁸⁶Sr/⁸⁸Sr = 0.11940.

* Krueger Enterprises Inc., Cambridge, Massachusetts 02139, USA.

† Conoco Inc., Ponca City, Oklahoma 74603, USA.

‡ Geochemical Institute, Göttingen, FRG.

The ¹⁸O depletion of oxygen in calcite 582 (Fig. 2, Table 1) is most unusual because calcite concentrates ¹⁸O (refs 12, 13). However, duplicate analyses in two different laboratories have confirmed that calcite 582 does, in fact, have $\delta^{18}\text{O}$ (SMOW) values between -19.8 and -21.1%. Additional analyses of three other specimens of calcite indicate that their $\delta^{18}\text{O}$ (SMOW) values range from -12.91% to -14.2%. They are therefore also anomalously depleted in ¹⁸O, although not as much as sample 582.

A plausible explanation for the unusual oxygen-isotope composition of the calcites is that they precipitated from glacial meltwater which is known to be highly depleted in ¹⁸O relative to sea water⁸. For example, the $\delta^{18}\text{O}$ (SMOW) values of ice exposed at the Elephant Moraine range from -42.6% to -49.7%

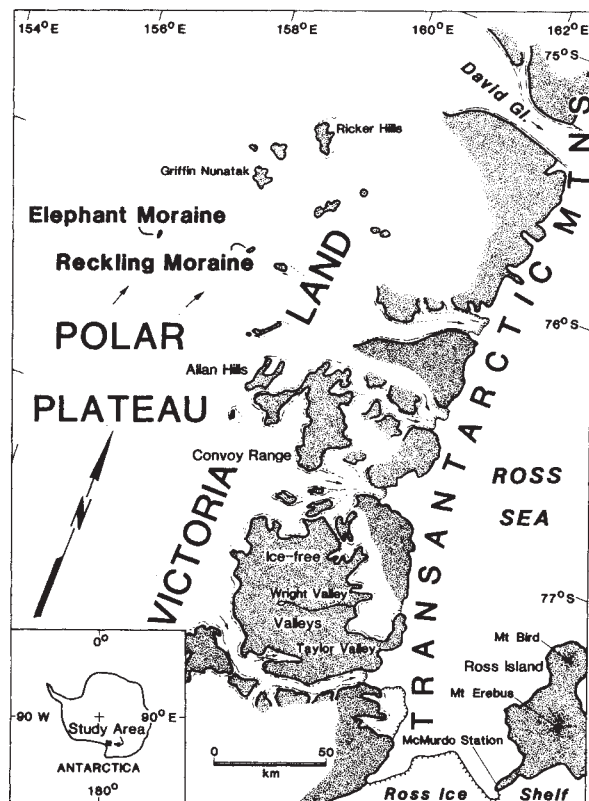


Fig. 1 Map of Victoria Land and the edge of the East Antarctic ice sheet showing the locations of the Elephant and Reckling Moraines.

with an average of $-47.4 \pm 0.6\%$ (one standard deviation of the mean)⁶.

When calcite precipitates from an aqueous solution under equilibrium conditions, the isotope fractionation factor (α) for oxygen is temperature dependent according to the equation⁹:

$$1,000 \ln \alpha_w^c = 2.78 \times 10^6 T^{-2} - 2.89$$

Therefore, the calcite in sample 582 could have precipitated at a temperature of $\sim 28^\circ\text{C}$ from water having $\delta^{18}\text{O}$ (SMOW) = -47.7% formed by melting of ice underlying the Elephant Moraine. The approximate oxygen-isotope equilibration temperatures of the other samples of calcite in Table 1 are close to 0°C . If the meltwater had been enriched in ^{18}O by isotope exchange with oxygen in silicate or carbonate minerals of the underlying bedrock, the estimated temperatures would increase by $\sim 5^\circ\text{C}$ for each 1% enrichment of the water in ^{18}O . Indirect evidence that the water was, in fact, somewhat more enriched in ^{18}O than glacial meltwater from the Elephant Moraine is provided by the oxygen-isotope composition of chert in sample EMA-2 (Table 1).

The oxygen in this chert ($\delta^{18}\text{O} = +13.03\%$, SMOW) did not equilibrate isotopically with glacial meltwater from the Elephant Moraine because the required equilibration temperature is unreasonably low (-35°C)⁹. However, when the temperature of deposition is arbitrarily set at 5°C , the $\delta^{18}\text{O}$ value of the water is found to be -30.3% (SMOW). Water of such isotope composition may result by mixing of glacial meltwater of the East Antarctic ice sheet (-47.4%) with groundwater that had previously exchanged oxygen with rock-forming silicate or carbonate minerals, all of which are enriched in ^{18}O relative to meteoric water.

Therefore, the isotopic compositions of oxygen in the calcites and the chert from the Elephant Moraine suggest that these minerals formed from water which was highly depleted in ^{18}O

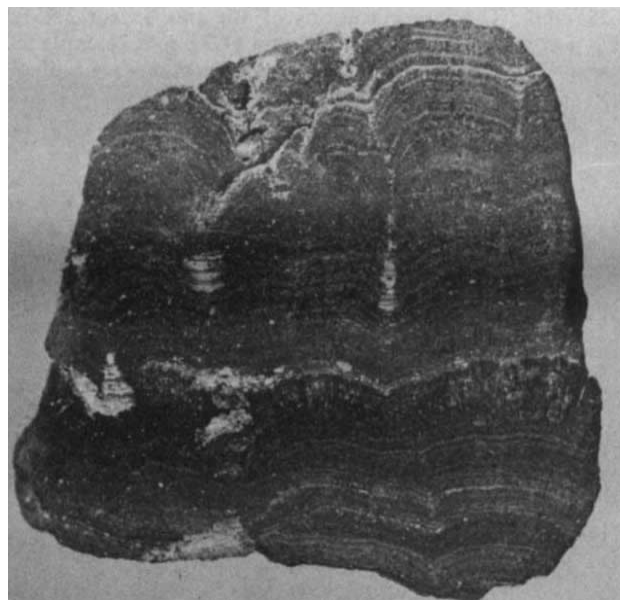


Fig. 2 Photograph of black calcite (sample 582, Table 1) from the Elephant Moraine, Antarctica. The specimen is 11 cm in diameter.

to an extent seen only in meteoric precipitation in the polar regions of the Earth.

The $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of five specimens of the black calcite in Table 1 are strongly clustered around a mean of 0.71435 ± 0.00022 (one standard deviation of the mean). This value exceeds the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of marine Sr throughout geological time and therefore rules out a marine origin for the calcite^{10,5}. However, calcite cleats in coal seams in the Transantarctic Mountains of southern Victoria Land have similar $^{87}\text{Sr}/^{86}\text{Sr}$ ratios (0.71233 to 0.72392)¹¹. The high $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of the calcite cleats indicate that the Sr could have originated from detrital feldspar and other Rb-bearing minerals in the sandstone of the Beacon Supergroup.

The $\delta^{13}\text{C}$ values of the calcites in Table 1 range within narrow limits from -22.0% (F-84-6A) to -23.1% (583) indicating strong depletion in ^{13}C . Only calcites containing carbon of biogenic origin derived by oxidation of plant remains or fossil fuels (coal, petroleum, natural gas) have such isotopic compositions. Coal in the Transantarctic Mountains has $\delta^{13}\text{C}$ values ranging from -23.5% to -26.0% ¹¹. The $\delta^{13}\text{C}$ values of calcite cleats within these coal seams range from -15.6% to -16.9% indicating that a major proportion of the carbon is of biogenic origin¹. A comparison of $\delta^{13}\text{C}$ values therefore supports the conclusion that the carbon in the black calcite clasts on the Elephant Moraine originated in large measure from coal and/or abundant coaly plant fragments in the sedimentary rock formations of the Beacon Supergroup. Lumps of coal actually are a minor but common constituent of the Elephant Moraine⁴.

The total organic carbon concentrations of the calcites are only 0.30% (584) and 0.15% (582). Pyrolysis of the organic carbon indicates that sample 584 contains 0.73 mg g^{-1} distillable hydrocarbons and 0.85 mg g^{-1} of hydrocarbons derivable by thermal decomposition of kerogen in the sample. The low organic carbon content indicates that the colour of the calcite crystals is probably caused by small amounts of Fe, Mn and Mg rather than by the presence of organic carbon¹.

The average concentrations of trace metals in five samples of calcite from the Elephant Moraine are generally low: Ag = 0.16 p.p.m. ; As = 1.3 p.p.m. ; Au = 0.012 p.p.m. ; Cu = 4.7 p.p.m. ; Mo = 0.87 p.p.m. ; Pb = 1.5 p.p.m. ; Zn = 2.8 p.p.m. and Se < 2.0 p.p.m. In addition, the concentrations of Hg, Tl, Bi, Cd, Ga, Pt, Sn and Te are $<0.5 \text{ p.p.m.}$ and those of Sb and Pd are

<0.25 p.p.m. The concentrations of Rb and Sr are 3.75 and 63.6 p.p.m. (sample 582) and 3.83 and 11.2 p.p.m. (sample 583), respectively. Such low trace metal concentrations are typical of calcite that precipitated from aqueous solutions at low temperature and therefore do not suggest that potentially valuable mineral deposits formed from the subglacial hot springs.

The unusual depletion of the calcites and chert in ^{18}O is compelling evidence that both minerals formed in contact with glacial meltwater of the East Antarctic ice sheet. The deposit from which the clasts were derived could not have formed at a time when the ice sheet had retreated from the Transantarctic Mountains because, the last time that happened during the Pliocene epoch before ~2.5 Myr, a marine embayment occupied the area west of the Transantarctic Mountains¹⁴. We conclude, therefore, that calcite and chert were precipitated by springs that discharged water formed in part by melting of the present ice sheet. The necessary heat required for melting may have been provided by a geothermal anomaly under the ice related to the volcanic activity on Ross Island and in Victoria Land which started at ~18 Myr and has continued to the present¹⁵. The principal volcanic centres lie along major lineaments that strike north-south and may represent deep faults along the front of the Transantarctic Mountains¹⁷. Mount Erebus, the only currently active volcano in the area, lies at the intersection of east-west and north-south trending lineaments^{16,17}. The most recent volcanic activity in southern Victoria Land was dated at 0.08 ± 0.13 Myr^{15,18}. Therefore, geothermal anomalies could have occurred under the ice sheet west of the Transantarctic Mountains within the past 2.5 Myr.

The development of hot spots under the ice sheet may have caused convection of solute-rich groundwater augmented by melting at the base of the ice sheet. The resulting springs discharged aqueous solutions of varying isotopic compositions and temperatures that precipitated calcite and amorphous silica. The calcite formed mounds of calcareous sinter or travertine¹⁹. The elevated $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of the calcites indicate that the solutions contained strontium derived from old Rb-bearing minerals in the underlying bedrock. In addition, the solutions contained bicarbonate ions derived from biogenic carbon that originated from coal seams or coaly plant fragments in the sandstones of the Beacon Supergroup. The deposits built up by the subglacial springs were eventually overridden by the ice sheet and fragments were transported to the Elephant Moraine which is forming at the present time by ablation of basal ice that is forced to the surface by a subglacial bedrock ridge⁴.

We thank W. A. Cassidy and John Schutt for the calcite samples they donated. David Buchanan and Karen S. Taylor assisted with field work on the Elephant Moraine during the 1986-87 and 1985-86 seasons, respectively. We thank Conoco, Inc. and Exlog/Brown and Ruth Laboratories for the use of laboratory facilities. Robert Tope drafted Figure 1. This research was supported by the NSF.

Received 14 September 1987; accepted 25 January 1988.

1. Sipiera, P. P. & Landis, C. A. *Lunar planet. Inst. tech. Repl.* 86-01, 101-103 (1986).
2. Cassidy, W. A. & Schutt, J. W. *Antarct. J. U.S.* 19, 39-40 (1984).
3. Cassidy, W. A. *et al. Antarct. J. U.S.* 18, 80-82 (1983).
4. Faure, G. & Taylor, J. *U.S.* 20, No. 5, 11-12 (1985).
5. Faure, G. *et al. Antarct. J. U.S.* 21 (in the press).
6. Faure, G. *et al. Antarct. J. U.S.* 22 (in the press).
7. Read, J. F. in *Developments in Sedimentology 20, Stromatolites* (ed. Walter, M. R.) (Elsevier, Amsterdam, 1976).
8. Lorus, C. *et al. Nature* 316, 591-596 (1985).
9. Friedman, I. & O'Neil, J. R. in *Data of Geochemistry* 6th edn (ed. Fleischer, M.) (US Geological Survey, Washington, 1977).
10. Burke, W. H. *et al. Geology* 10, 516-519 (1982).
11. Faure, G. & Botoman, G. *Chem. Geol.* 2, 313-324 (1984).
12. Faure, G. *Principles of Isotope Geology* 2nd edn (Wiley, New York, 1986).
13. Hoefs, J. *Stable Isotope Geochemistry* 3rd edn (Springer, Heidelberg, 1987).
14. Webb, P. N. *et al. Geology* 12, 287-291 (1984).
15. Armstrong, R. L. *N.Z. J. Geol. Geophys.* 21, 685-698 (1978).
16. McIver, J. R. & Gevers, T. W. *Trans. geol. Soc. S. Afr.* 73, 65-88 (1970).
17. Kyle, P. R. & Cole, J. S. *Bull. Volcanol.* 38, 16-25 (1974).
18. Armstrong, R. L. *et al. Science* 159, 187-188 (1968).
19. Bates, R. L. & Jackson, J. A. *Glossary of Geology* 2nd edn (Am. Geol. Inst., Falls Church, Virginia, 1980).

Semiochemical basis of the retinue response to queen honey bees

Keith N. Slessor*†, Lori-Ann Kaminski§, G. G. S. King†, John H. Borden§ & Mark L. Winston§

† Department of Chemistry and the §Center for Pest Management, Department of Biological Science, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

* To whom correspondence should be addressed

The substance produced by the mandibular gland complex (HQMC) of queen honey bees is considered to be responsible for retinue formation in honey bees, *Apis mellifera* L.¹⁻³. The retinue response includes the licking and antennating behaviour which signals the presence of a dominant reproductive queen and thereby establishes and stabilizes the social fabric of the colony. Despite 25 years of research, none of the known constituents of HQMC extract^{4,5} has been effective in eliciting full retinue behaviour. We report that the mandibular-gland-based retinue response is mediated by five semiochemicals. Each component is weakly active alone, but the complete blend imparts activity equivalent to HQMC extract at a level as low as 10^{-7} of that present in a queen.

Laying queens were immobilized at -20 °C. Mandibular glands were excised, macerated and extracted twice in HPLC-grade methanol to give a combined total extract of 100 µl per queen. A 3-queen equivalent (Qeq) HQMC extract was chromatographed on a micro-column of acid-treated silica gel. The column was eluted with mixtures of heptane and ether to give ten fractions which were analysed by splitless capillary gas

Table 1 Response of worker honey bees to glass pseudo-queens treated with HQMC components alone and in all combinations

Treatments	Mean ($\pm$ s.e.m.) number of bees contacting stimulus
Queen extract (10^{-3} Qeq)	24.7 $\pm$ 3.3. a
HOB+9ODA+9HDA+HVA	17.1 $\pm$ 3.4 ab
HOB+9ODA+HVA	13.1 $\pm$ 2.3 bc
HOB+9HDA+HVA	6.1 $\pm$ 1.3 cd
HOB+9ODA+9HDA	6.7 $\pm$ 1.7 cd
9ODA+9HDA+HVA	3.4 $\pm$ 0.6 d
HOB+HVA	7.6 $\pm$ 1.1 cd
HOB+9ODA	9.0 $\pm$ 1.8 cd
9ODA+HVA	4.3 $\pm$ 1.1 d
9ODA+9HDA	3.5 $\pm$ 1.0 d
HOB+9HDA	3.5 $\pm$ 0.6 d
9HDA+HVA	3.9 $\pm$ 0.7 d
9ODA	6.1 $\pm$ 1.3 cd
HVA	4.7 $\pm$ 0.9 d
9HDA (76% R-(-))	3.5 $\pm$ 1.3 d
HOB	2.9 $\pm$ 0.7 d
Methanol (solvent blank)	1.3 $\pm$ 0.4 d

Glass pseudo-queens were fashioned from a Pasteur pipette (2.5 cm of the top and 5 cm of the narrowed bottom) cut and sealed at both ends, with a small indentation for receiving a 10 µl test stimulus in the 'head' end. Bioassay arenas each containing 15 worker bees were prepared from plastic Petri dishes (15 $\times$ 2 cm). A pseudo-queen was inserted through a hole cut in the side, and activity within the cage was videotaped for 5 min. At 30 s intervals the number of workers contacting the pseudo-queen were recorded. These counts were summed for each replicate. All treatments containing HQMC components were presented at concentrations corresponding to 10^{-3} Qeq. The queen extract contained HOB (13 µg) + 9ODA (150 µg) + 9HDA (71% R-(-))(55 µg) + HVA (1.5 µg) per queen, arbitrarily defined as 1 Qeq, and was prepared from a mixture of eight mated queens, three of which were laying, and five of which were received in mailing cages. Means followed by the same letter were not significantly different, Tukey's multiple range test²⁰, $P < 0.05$, $n = 15$.

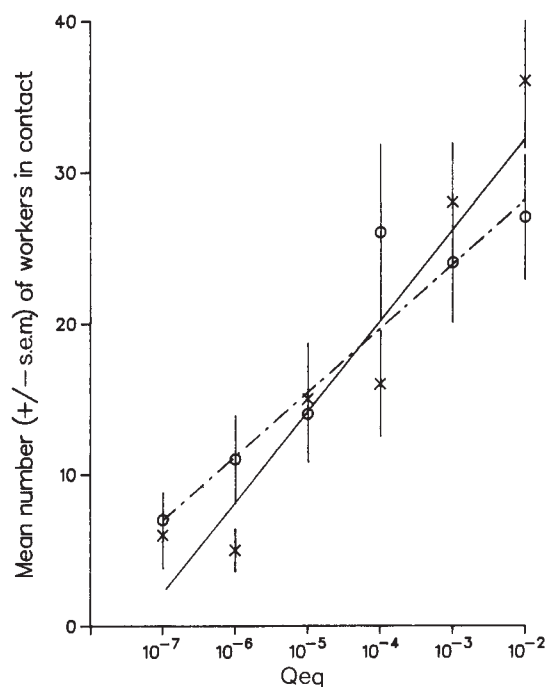


Fig. 1 Relationship between dose and contact response in the pseudo-queen bioassay for natural (crosses) and synthetic (circles) HQMC. See Table 1 for definition of 1 Qeq. All data adjusted to eliminate solvent response. Regression equations and significance for natural and synthetic HQMC, respectively, are $y = 46.0 + 6.29 \log x$, $r^2 = 0.90$, $P < 0.005$ and $y = 37.6 + 4.31 \log x$, $r^2 = 0.86$, $P < 0.005$ (ref. 20).

chromatography (SCGC) before and after methylation with BF_3 -methanol⁶.

In micro-pipette bioassays (L.-A.K., K.N.S. and M.L.W., manuscript in preparation), the ten fractions (F) were analysed alone and in all possible binary combinations. The most striking feature of the resulting data matrix was the broad behavioural activity elicited by the binary combinations. When F3 or F4 was combined with F8 the response by worker bees was consistently high. The most abundant substance in F3 and F4 was destroyed by acetylation, resulting in a substance shown by SCGC-mass spectroscopy (MS) to be methyl *p*-acetoxybenzoate. The parent material, methyl *p*-hydroxybenzoate (HOB), is known in HQMC^{7,8}. Bee responses to HOB alone were 9% of those for HQMC ($n = 5$).

Weakly active F6 (22% of HQMC) contained 9-keto-2(*E*)-decanoic acid (9ODA), also a known queen pheromone⁹⁻¹¹. A stronger response (52% of HQMC) occurred to HOB when combined with 9ODA, in agreement with the binary fraction bioassays and a report that these chemicals can attract swarms¹¹.

9-Hydroxy-2(*E*)-decanoic acid (9HDA) was present in the weakly active F7 (20% of HQMC). The presence of 9HDA in HQMC is known¹² and the predominant enantiomer, *R*-(-)-9HDA, can maintain queenless swarms¹³. Response to racemic 9HDA when combined with 9ODA was minimal (4% of HQMC), but when it was combined with HOB, activity increased to 43% of HQMC. When HOB, 9ODA, and racemic 9HDA were combined, activity rose to 57% of HQMC.

In F8, there was a final, more polar substance, which when acetylated and analysed by SCGC-MS gave a spectrum reminiscent of 4-hydroxy-2-methoxyphenylethanol¹⁴. Synthesis of this compound¹⁵ led to two hydroxymethoxyphenylethanol, the acetate derivatives of which had SCGC-retention characteristics different from the F8 component. Neither was active when tested with the other materials. Therefore, the active isomer (2-3 μg) was isolated and identified from a HQMC extract from 5 queens (these queens were extracted in autumn when pheromone con-

Table 2 Retinue response by worker honey bees on hive frames to the colony's own queen and pseudo-queens treated with HQMC extract, synthetic HQMC and components

Treatments	Mean ($\pm$ s.e.m.) number of bees contacting stimulus
Colony's own tethered queen	86.5 $\pm$ 2.8 a
Pseudo-queen with HQMC extract	51.6 $\pm$ 2.6 b
Pseudo-queen with synthetic HQMC	45.1 $\pm$ 3.7 b
Pseudo-queen with 9ODA + (76% <i>R</i> -(-)-9HDA	12.1 $\pm$ 2.9 c
Pseudo-queen with solvent	4.3 $\pm$ 1.3 c

All treatments containing HQMC components were presented at concentrations corresponding to 10^{-2} Qeq, see Table 1 for composition. Means followed by the same letter were not significantly different, Tukey's multiple range test²⁰, $P < 0.05$, $n = 10$.

tent is low). High resolution ^1H nuclear magnetic resonance spectroscopy gave a spectrum interpreted as the diacetate of 4-hydroxy-3-methoxyphenylethanol (HVA). Acetylation of HVA provided NMR, SCGC-MS spectra and retention times identical to those obtained from the acetylated derivative in F8.

In a new pseudo-queen bioassay (Table 1), a combination of the five compounds in an approximately natural ratio induced behaviour equivalent to that evoked by HQMC extract. Quantities of mandibular components were established by SCGC on trimethylsilylated derivatives measured by a flame ionization detector calibrated with known standards.

Three experiments verified that the identified compounds reproduced the retinue response induced by HQMC extract. Over a broad concentration range, responses to the HQMC extract and the synthesis mixture were not statistically different¹⁶ in slope ($P > 0.2$) and intercept ($P > 0.2$) (Fig. 1). Synthetic queen substance was active at or below 10^{-7} Qeq (analysis of variance, $P < 0.025$, $n = 16$). Secondly, when the enantiomers of 9HDA were tested alone, together and in combination with the other mandibular components, both enantiomers were essential for full retinue behaviour ($P < 0.05$, $n = 10$), the fourth known example of enantiomeric synergism between insect semiochemicals¹⁷⁻¹⁹.

To ensure that the pseudo-queen bioassay was a valid indicator of natural retinue response, a third experiment was conducted on occupied frames. Five frames containing roughly equal numbers of bees were taken from a colony from which the queen had been removed for at least an hour. These frames were placed on a viewing rack on a sunny, warm (23 $^\circ\text{C}$) day. Five treatments were randomly assigned and each placed in the centre of a frame. The treatment area was video-taped and bee behaviour analysed as for the pseudo-queen bioassay. New sets of five frames, each set from a separate colony, were used for additional replicates. The colony's own queen produced the greatest response (Table 2), indicating the presence of non-chemical behavioural cues and/or additional pheromones not found in the HQMC. Synthetic and natural mandibular extracts were equally active, once again showing that we have accounted for all of the activity in HQMC.

In addition to recognition of the queen, the five mandibular gland semiochemicals may also be involved in worker orientation during swarming, suppression of worker ovary development and egg laying, inhibition of queen rearing, individual queen recognition, and other queen-controlled activities. We are currently re-investigating these functions. In addition we are exploring the potential benefits to be derived from an improved ability to manage and control this most complex of social insects.

We thank C. D. Andrews, G. G. Grant, R. C. Hutchinson, M. Juriga, S. A. Kolmes, X.-Z. Meng, G. Owen, R. D. Routledge, A. S. Tracey, M. M. Tracey and M. H. Wyborn for assistance, and the Natural Sciences and Engineering Research Council of Canada and the Wright Institute for financial support.

Received 9 November 1987; accepted 2 February 1988.

1. Gary, N. E. *Science* **133**, 1479–1480 (1961).
2. Velthuis, H. H. W. *Behavior* **41**, 105–129 (1972).
3. Butler, C. G., Callow, R. K., Koster, C. G. & Simpson, J. J. *Apic. Res.* **12**, 159–166 (1973).
4. Callow, R. K., Chapman, J. R. & Paton, P. N. *J. Apic. Res.* **3**, 77–89 (1964).
5. Simpson, J. J. *Apic. Res.* **18**, 233–249 (1979).
6. Slessor, K. N., King, G. G. S., Miller, D. R., Winston, M. L. & Cutforth, T. L. *J. chem. Ecol.* **11**, 1659–1667 (1985).
7. Pain, J., Hugel, M.-F. & Barbier, M. C. *r. Acad. Sci. Paris* **251**, 1046–1048 (1960).
8. Barbier, M., Lederer, E., Reichstein, T. & Schindler, O. *Helv. chim. Acta* **43**, 1682–1689 (1960).
9. Barbier, M. & Lederer, E. C. *r. Acad. Sci. Paris* **250**, 4467–4469 (1960).
10. Callow, R. K. & Johnston, N. C. *Bee World* **41**, 152–153 (1960).
11. Silverstein, R. M. & Young, J. C. in *Pest Management with Insect Sex Attractants* (ed. Beroza, M.) 1–29 (Am. Chem. Soc., Washington, 1976).
12. Butler, C. G. & Faurey, E. M. *J. Apic. Res.* **3**, 65–76 (1964).
13. Winston, M. L., Slessor, K. N., Smirle, M. J. & Kandil, A. A. *J. chem. Ecol.* **8**, 1283–1288 (1982).
14. Crewe, R. M. & Velthuis, H. H. W. *Naturwissenschaften* **67**, 467–469 (1980).
15. Howard, A. S., Michael, J. P. & Schmidt, M. A. *S. Afr. Tydskr. Chem.* **34**, 132–133 (1981).
16. Kleinbaum, D. G. & Kupper, L. L. *Applied Regression Analysis and Other Multivariable Methods* (Duxbury, North Scituate, 1978).
17. Borden, J. H., Chong, L., McLean, J. A., Slessor, K. N. & Mori, K. *Science* **192**, 894–896 (1976).
18. Lanier, G. N., Classon, A., Stewart, T., Piston, J. J. & Silverstein, R. M. *J. chem. Ecol.* **6**, 677–687 (1980).
19. Millar, J. G., Pierce, H. D. Jr, Pierce, A. M., Oehlschlager, A. C. & Borden, J. H. *J. chem. Ecol.* **8**, 1071–1081 (1985).
20. Zar, J. H. *Biostatistical Analysis* 2nd edn (Prentice-Hall, Englewood Cliffs NJ 1984).

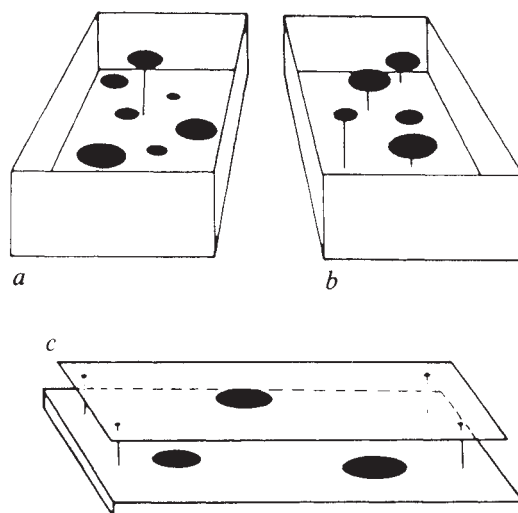


Fig. 1 Design of artificial meadow and flowers in training conditions (a) and test conditions (b). The floor of the meadow was 30×40 cm and the height of the wall was 12 cm. Flowers were cardboard disks ranging between 20 mm and 70 mm in diameter. The height of the flower carrying the reward in training conditions was 70 mm. No reward was offered during tests. In one experiment, flowers were black on white meadow; in two other experiments, flowers were blue on yellow meadow, using two different colour combinations (see Fig. 2 legend). c, Modified set-up using a perspex sheet placed 5 cm above the meadow to restrict the bees' flight domain. The meadow was white, surrounded by a 30 cm high wall (not shown). Black flowers were used, chosen randomly three at a time from a pool of seven flowers ranging between 25 mm and 100 mm in diameter. The bees were trained (see text) to fly to a flower that was either the highest (as shown), the lowest, or intermediate in height (in which case a second perspex sheet 5 cm above the first was used). As before, no reward was offered during tests. Here, as well as in (a) and (b), sizes and positions of all flowers were varied randomly between training runs and between individual tests. Olfactory cues were excluded by changing the flower carrying the reward frequently, and by using new flowers (or a fresh perspex sheet) during tests.

Motion cues provide the bee's visual world with a third dimension

M. Lehrer*, M. V. Srinivasan, S. W. Zhang* & G. A. Horridge

Centre for Visual Sciences, Research School of Biological Sciences, Australian National University, PO Box 475, Canberra, ACT 2601, Australia

To extract the third dimension from a two-dimensional retinal image most insects, including bees, cannot rely on mechanisms common in vertebrates such as accommodation, binocular convergence or stereoscopic vision^{1,2}. Instead, they use the apparent size of familiar objects (the nearer the object, the larger its image), and objects' apparent motion (the nearer an object, the higher the speed of its image)^{3–8}. In several studies^{9–12} bees have been found to exploit size cues, whereas in others^{6,11,13} they seem to use both strategies. We have studied the influence of motion cues in isolation by excluding size cues. We report that bees can discriminate between objects at different distances irrespective of their size. This discrimination is mediated primarily by the green-sensitive visual channel and is therefore colour blind, like all of the motion-dependent behaviours investigated so far in the bee^{14–17}. The bee's ability to discriminate range by motion of the image explains how bees manage to manoeuvre in novel environments, where the size of objects is unknown.

We trained bees to collect food on a white 'meadow' surrounded by a white wall (Fig. 1a). Six black 'flowers' were placed on the floor of the meadow, and one was raised on a thin 70 mm-high 'stalk' and provided with a reward of sugar solution. Each rewarded visit of a bee to the meadow, termed "training run", was recorded. The sizes and positions of all flowers were varied randomly between training runs, except that the flower carrying the reward was always 70 mm high. This was to ensure that the bees were trained to associate only the height of a flower with the reward, and not its angular subtense, its position with relation to the other flowers or other possible visual references.

During tests the bees were offered five flowers, of different sizes, each placed at a different height, the highest being at the training height and the lowest on the floor (Fig. 1b). None of

the flowers in a test carried a reward. In these tests, the distribution of the bees' landings on the five flowers (Fig. 2a) was strictly correlated with flower height ($r = 0.97$, $P < 0.01$). Thus, the bees can estimate the relative heights of the flowers irrespective of their apparent size.

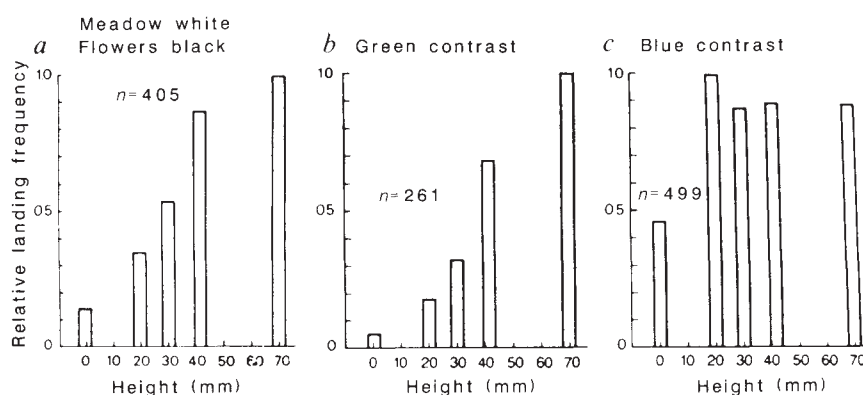
Video recordings revealed that the searching bees flew primarily in straight lines above the flowers. The resulting flow field would be rich in translatory components and therefore in depth-related cues.

In the next two experiments the meadow was yellow and the flowers blue. We selected two colour combinations (see legend to Fig. 2a and b). In one experiment, contrast between flowers and background was restricted to the green receptors, in the other to the blue receptors. (For colour contrast calculations see refs 15 and 17.)

With green contrast (Fig. 2b), after 30 training runs, the bees' ability to discriminate between flowers was equal to that in the first experiment with the black and white colour combination (Fig. 2a). In the absence of green contrast, however, the bees' visual discrimination was very poor even after 80 training runs (Fig. 2c). We conclude that the bees' performance in our experiments was mediated chiefly by the green-sensitive visual channel and is consequently colour blind. This strongly suggests that the cues involved are indeed motion cues because several types of motion-dependent behaviour in the honey bee have been shown to be green sensitive and colour blind. These include: the optomotor response¹⁴, the movement-avoidance response¹⁵, scanning behaviour¹⁶, high spatial acuity during flight¹⁷, and spatial localization (M.L., manuscript in preparation).

* Present addresses: Zoologisches Institut der Universität Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland (M.L.); Institute of Biophysics, Academia Sinica, Beijing, China (S.W.Z.)

Fig. 2 Distribution of the bees' landings on five flowers at different heights, after having trained the bees to the highest flower. Landing frequencies are normalized to a maximum value of 1. n = Total number of landings. *a*, Black flowers on a white meadow, resulting in a contrast of 80% to both the blue and the green receptors. Tests began after 30 training runs. *b*, Blue flowers (pigment-paper CV 122, Colour VU Productions, New York) on a yellow meadow (CV 5). Contrast (40%) was restricted to the green receptors. Tests began after 30 training runs. *c*, Blue flowers (CV 122) on a yellow meadow (CV 7), resulting in a contrast (56%) restricted to the blue receptors. In this case, tests began after 80 training runs.



To examine the possibility that the bees' preference for the highest flower (Fig. 2*a* and 2*b*) was simply due to an innate tendency to select the fastest moving object¹⁸, and was not a consequence of training, we placed a transparent perspex sheet 5 cm above the floor (Fig. 1*c*), thus preventing the bees from flying in plains lower than this height. This allowed us to train bees either to the fastest moving flower, or to a slower moving one. Three black flowers were used, each of which could be placed either 'low' (on the floor) or 'high' (affixed to the underside of the perspex sheet). One flower, either a high or a low one, had a sugar-water reward placed on top of the perspex sheet inside its area of projection. The two other flowers were placed at the alternative height (Fig. 1*c*). Sizes and positions of the flowers were varied continually, as before. In subsequent tests, begun after 30 training runs, we recorded the bees' landings on the area of projection of each flower on the perspex sheet above it.

Bees trained to expect a reward on a high flower strongly preferred the high flower in the tests; the landing frequency on the high flower was 73% ($n=488$), compared with an expected frequency of 33.3% if choice was random. Similarly bees trained to expect a reward on a low flower subsequently preferred the low flower, landing on it in 84.4% of the cases ($n=243$). In a third experiment we used two perspex sheets, one 5 cm and the other 10 cm above the floor. One flower was placed on the floor, one on the lower perspex sheet, and one affixed to the underside of the higher sheet. The reward was placed above the intermediate-level flower, and this received 68.6% of the bees' landings ($n=461$) during subsequent tests.

These results demonstrate that bees can be trained to select objects by their distance, irrespective of the objects' sizes. Further experiments have shown that the bees ignored not only the flowers' size, but also their shape in this task (manuscript in preparation). The conclusion that the bees used the motion of an image to discriminate distances in our experiments is based both on the exclusion of all other possible cues, and on the finding that this ability is impaired in the absence of green contrast. Our results raise further questions, including the role of different types of motion at the edges of objects which move against a textured background, and the participation of motion cues in pattern discrimination.

Thanks are due to Rüdiger Wehner for his comments on the manuscript, to Wolfgang Kirchner for stimulating discussions and help during tests, to Mr R. Dencio for assistance with the video-film analysis, and to Mr A. Saeck for providing and maintaining the bee colonies. M.L. and S.W.Z. extend their gratitude to Adrian Horridge for supporting their visiting fellowship at the Centre for Visual Sciences in Canberra.

- Wallace, G. K. *J. exp. Biol.* **36**, 512-525 (1959).
- Collett, T. S. *J. exp. Biol.* **76**, 237-241 (1978).
- Cartwright, B. A. & Collett, T. S. *J. exp. Biol.* **82**, 367-372 (1979).
- Erikson, E. S. *J. exp. Biol.* **86**, 337-340 (1980).
- Horridge, G. A. *Proc. R. Soc. Lond.* **B229**, 23-27 (1986).
- Lauer, J. & Lindauer, M. *Informationsaufnahme und Informationsbearbeitung im lebenden Organismus* Vol. 1, 1-87 (Steiner, Wiesbaden, 1971).
- Maurer, B. thesis, Univ. Zurich (1980).
- Lehrer, M. thesis, Univ. Zurich (1980).
- Cartwright, B. A. & Collett, T. S. *J. comp. Physiol.* **151**, 521-543 (1983).
- Cheng, K., Collett, T. S., Pickhard, A. & Wehner, R. *J. comp. Physiol.* **161**, 469-475 (1987).
- Kaiser, W. & Liske, E. *J. comp. Physiol.* **89**, 391-408 (1974).
- Srinivasan, M. V. & Lehrer, M. *J. comp. Physiol.* **155**, 297-312 (1984).
- Lehrer, M., Wehner, R. & Srinivasan, M. V. *J. comp. Physiol.* **157**, 405-415 (1985).
- Srinivasan, M. V. & Lehrer, M. *J. comp. Physiol.* **162**, 159-172 (1988).
- Erber, J. *J. comp. Physiol.* **146**, 273-282 (1982).

Population coding of saccadic eye movements by neurons in the superior colliculus

Choongkil Lee*, William H. Rohrer* & David L. Sparks†

Neurobiology Research Center, Department of Physiology and Biophysics, University of Alabama, Birmingham, Alabama 35294, USA

The deeper layers of the superior colliculus are involved in the initiation and execution of saccadic (high velocity) eye movements¹. A large population of coarsely tuned collicular neurons is active before each saccade. The mechanisms by which the signals that precisely control the direction and amplitude of a saccade are extracted from the activity of the population are unknown. It has been assumed²⁻⁶ that the exact trajectory of a saccade is determined by the activity of the entire population and that information is not extracted from only the most active cells in the population at a subsequent stage of neural processing. The trajectory of a saccade could be based on vector summation of the movement tendencies provided by each member of the population of active neurons⁴ or be determined by a weighted average of the vector contributions of each neuron in the active population². Here we present the results of experiments in which a small subset of the active population was reversibly deactivated with lidocaine. These results are consistent with the predictions of the latter population-averaging hypothesis and support the general idea that the direction, amplitude and velocity of saccadic eye movements are based on the responses of the entire population of cells active before a saccadic eye movement.

Neurons in the superior colliculus displaying saccade-related spike activity have movement fields, that is, each cell discharges

Received 14 December 1987; accepted 12 February 1988.

- Horridge, G. A. *Endeavour* **1**, 7-17 (1977).
- Collett, T. S. & Harkness, L. I. K. in *Analysis of Visual Behaviour* (eds Ingle, D. J., Goodale, M. A. & Mansfield, R. J. W.) 111-176 (MIT Press, Cambridge, 1982).
- Gibson, E. J., Gibson, J. J. & Smith, O. W. *J. exp. Psychol.* **58**, 40-51 (1959).

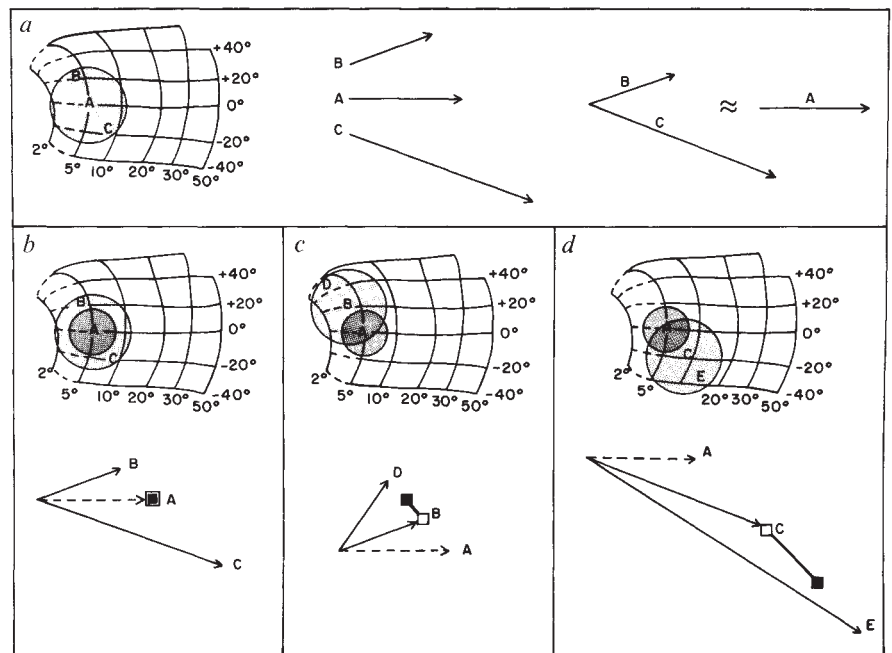
* Present addresses: Department of Psychology, Seoul National University, Seoul, Korea (C.L.); Department of Anatomy, University of Medicine and Dentistry of New Jersey, PO Box 55, Piscataway, New Jersey 08854, USA (W.H.R.).

† To whom correspondence should be addressed.

Fig. 1 *a*, The population-averaging scheme of Sparks *et al.*². Left: the motor map of the left superior colliculus based on Robinson's⁹ microstimulation study. Isoamplitude lines (2° to 50°) run from the lateral edge to the medial aspect of the superior colliculus. Isodirection lines (-40° to 40°) run from the rostro-lateral edge to the caudo-medial border of the colliculus. The stippled area represents the hypothetical extent of cells active before saccades to a target located 5° to the right of the fixation stimulus. The active population is assumed to be symmetrical in shape³. Moreover, it is known² that neurons in the middle of the active population fire most vigorously and that, for other cells, the peak firing rate decreases as a function of distance from the centre of the active region. Middle: cells at locations *A*, *B* and *C* fire most vigorously for the movements shown. Right: weighted averaging of activity at points *B* and *C* yields the same movement as activity at the centre of the active population (*A*). Note that a true average of vectors *B* and *C* does not yield vector *A* because of the non-linearity of the amplitude scale in the motor map. If one assumes a homogeneous distribution of cells and a symmetrical spatial distribution of firing rates (see above), a non-linear synaptic weighting of the contribution of cells at different locations to brainstem oculomotor nuclei is required by this scheme. *b-d*, The predicted effect of deactivating a subset of cells in the active population. The site of deactivation (darkly stippled circle) remains the same in each panel, but the location of the active population (lightly stippled area) is different in each panel because saccades to three different targets are required. Beneath each map are the saccade vectors associated with neural activity at each of the locations illustrated. The open square represents the vector of the intended, or programmed, saccade associated with activity in the lightly stippled area. The dashed line represents the vector of the movement tendency produced by neurons at the deactivated site. These neurons will not contribute to the metrics of the saccade, and a saccade to the approximate location of the filled square is predicted.

Methods. Three monkeys (*Macaca mulatta*) were trained to look to visual targets for a liquid reward. A laboratory computer controlled the experiments and compared eye and target positions. Eye position was monitored by the scleral search-coil technique¹⁵. On each trial, the animal fixated a centrally located visual stimulus for a variable period. Then the fixation stimulus was extinguished and an eccentric target appeared simultaneously. Reward was contingent on the occurrence of a saccade that directed gaze toward the eccentric stimulus. An insulated silver-coated glass pipette¹⁴ filled with 2% lidocaine hydrochloride was used for extracellular recording of unit activity, for the reversible inactivation of neural elements near the tip of the probe, and for electrical stimulation (stimulation trains 50 ms in duration at a rate of 500 Hz; each pulse was 0.2 ms; current ranged from 25–40 μ A). The location of the pipette tip in the collicular motor map was determined from movement field plots and measurements of the direction and amplitude of stimulation-induced saccades. The behavioural effects of lidocaine injections were assessed by comparing saccades to a selected set of targets before and after deactivation of collicular neurons.

in association with saccades having a particular range of directions and amplitudes^{2,7,8}. These neurons form a map of motor (saccadic) space as they are topographically organized on the basis of their movement fields^{2,7,9}. The high-frequency burst of spike activity generated by collicular neurons is tightly coupled temporally to saccade onset¹⁰, but spatially the movement fields are large and coarsely tuned². Consequently a large population of cells is active before each saccade. Here we reversibly deactivated a small subset of the active population with lidocaine. Figure 1 outlines the rationale of the experiment and illustrates the main predictions of the population-averaging hypothesis. It is assumed that the population of neurons discharging before a saccade occupies a symmetrical area³. For example, the hypothetical region of neurons active before a 5° rightward saccade is shown schematically by the stippled region in panel *a*. Neurons in the middle of this active population (*A*) discharge maximally before the intended saccade (a straight right saccade 5° in amplitude). Referring to the motor map, neurons located at point *B* discharge maximally before upward (20° angle) saccades 4° in amplitude and neurons at point *C* discharge maximally before downward (-20° angle) saccades 9° in amplitude. The population-averaging hypothesis suggests that for each subset of active neurons (*B*) tending to produce movements other than the programmed saccade, there will be another subset of equally active neurons (*C*) tending to produce a movement such that the weighted average effect of neural activity at points *B* and *C* will result in a saccade with the programmed direction and amplitude (panel *a*, right). According to this hypothesis, each member of the active population contributes to the ensuing



saccade. Detection of the locus of the most active elements in the population is unnecessary.

Suppose that a small region of collicular neurons (the darkly stippled area in Fig. 1, panel *b*), normally discharging maximally before straight right saccades 5° in amplitude, were deactivated with lidocaine. If the animal were required to make a 5° rightward movement, according to the population-averaging hypothesis, saccadic direction and amplitude would be unaffected (panel *b*). The weighted vector average of the movement tendencies produced by the unaffected active cells (the lightly stippled area) will result in a saccade having the correct direction and amplitude. For example, the weighted average of the movements produced by neurons near points *B* and *C* outside the deactivated region is still a 5° rightward movement. Referring to panel *c*, with the same region of neurons deactivated, if the animal is now required to make an upward (20° angle) saccade 4° in amplitude, the active population shifts to a location centred around point *B*. The deactivated neurons (normally discharging maximally before saccades that are larger and have less of an upward component than the programmed movement) will not counterbalance the effect of other cells (location *D*) firing maximally before saccades that are smaller and have more of an upward component than the intended movement. Consequently, the animal should make a hypometric saccade that has too much of an upward component (filled square). Similarly, with the same region of neurons deactivated, if the animal attempts to make a downward (-20° angle) 9° amplitude saccade, the resulting movement should be hypermetric and have too much of a downward component (panel

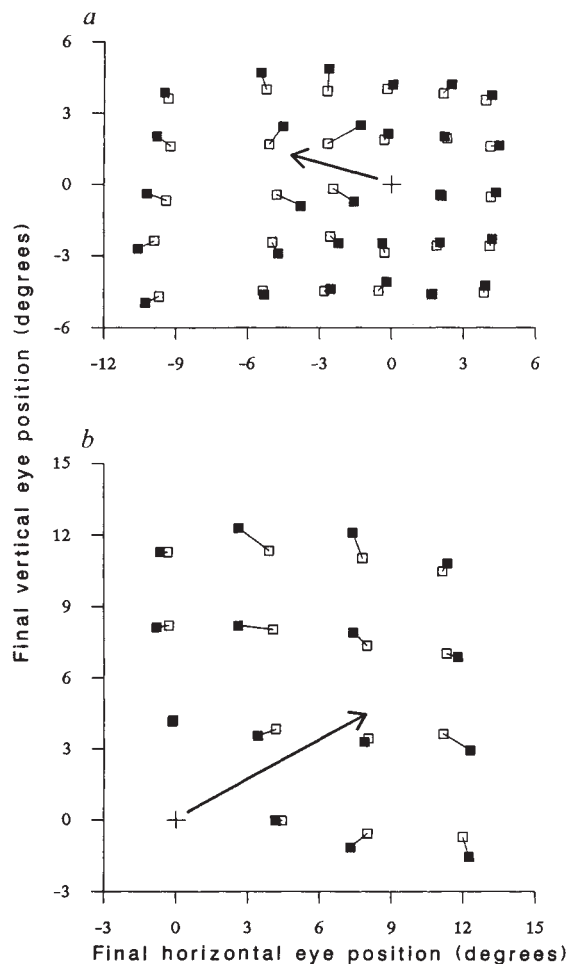


Fig. 2 Effects of a single 200 nl injection of lidocaine on the amplitude and direction of saccades to visual targets. The axes represent the horizontal and vertical end points of visually guided saccades. The location of the fixation target is indicated by the +. Electrical stimulation of the superior colliculus drove the eyes to the location marked by the tip of the arrow. Each open square represents the average of the end points of visually guided saccadic eye movements to one target before deactivation; filled squares represent the average end points of saccades to the same targets after deactivation. Lines connecting the squares represent the average error introduced by deactivation of part of the active population. *a*, Results from an injection site where the 'best saccade' was up and leftward. *b*, Results from an injection site in the left superior colliculus with an up and rightward 'best saccade'.

d). Thus, with neurons in a single region of the colliculus deactivated, a predictable pattern of errors should be observed for attempted saccades within the movement fields of the deactivated cells.

The predictions of the vector summation model⁴ differ from those illustrated in Fig. 1. Inactivation of a subset of the active population should always produce hypometric saccades (regardless of the location of the inactivated cells within the population) because the amplitude of a saccade is determined by summing the movement contributions of each cell in the active population.

We found that the trajectory of saccades to visual targets was modified for 5–20 min after injections of 50–200 nl 2% lidocaine hydrochloride. Results typical of those obtained from three monkeys and a total of 17 injection sites in the rostral superior colliculus are shown in Fig. 2. Effects of deactivating a region of the deeper layers of the superior colliculus upon the direction and amplitude of visually guided saccades are illustrated. The plots show the position of the initial fixation (+) and the endpoint (the tip of the arrow) of the 'best saccade,' defined as the movement produced by stimulation at the injection site. Each

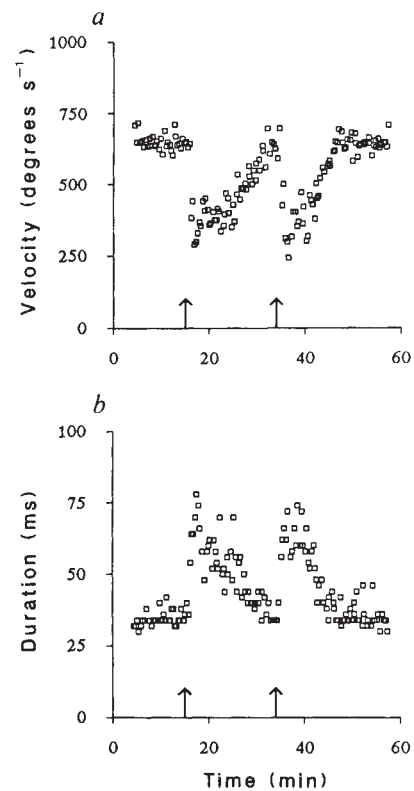


Fig. 3 The effects of two injections (each 100 nl 2% lidocaine) on visually guided saccades similar in direction and amplitude to the 'best saccade'. (The time of the injection is indicated by the arrows.) Each plotted point represents data from a single trial. For the data plotted, the saccade target was presented at the same location (6° to the right of and 10° above fixation). The reduction in vectorial velocity of $\sim 250^\circ \text{ s}^{-1}$ (*a*) was compensated for by an increase in saccadic duration (*b*).

symbol represents the average endpoint of 3 to 5 visually guided saccades. Unfilled symbols represent movements occurring before the injections; filled symbols represent post-injection trials for matching targets. Referring to Fig. 2*a*, note that, as predicted by the population-averaging hypothesis, saccades to targets requiring more of an upward component than the 'best saccade' had too much of an upward component, and saccades to targets requiring movements with more of a downward component than the 'best saccade' had too much of a downward component. Also as predicted, movements to targets requiring a saccade smaller in amplitude than the 'best saccade' were hypometric; saccades to targets requiring a movement larger than the 'best saccade' were hypometric. For the injection site represented in Fig. 2*a*, saccades to targets in the opposite half of the visual field were unaffected. The range of movements affected varied with the volume of the lidocaine injection.

For most injection sites, the direction and amplitude of movements similar to the 'best saccade' were not altered noticeably after the injection. Note that in Fig. 2*b*, saccades to targets requiring movements with a direction and amplitude similar to the 'best saccade' were quite accurate. But the velocity of movements similar to the 'best saccade' was dramatically reduced (Fig. 3*a*), although, unlike earlier reports^{11,12}, we found these reductions were accompanied by increases in duration (Fig. 3*b*), such that saccadic amplitude remained relatively constant.

The prediction of vector-summation models⁴ that inactivation of any subset of the active population invariably produces hypometric saccades was not confirmed. Instead, the pattern of errors in the amplitude (hypometria or hypermetria) and direction of visually guided saccades after deactivation of a localized region of collicular neurons was that predicted by the population-averaging hypothesis. In agreement with the suggestion that

the amount of activity of the active population is a determinant of saccadic velocity (ref. 13; D. P. Munoz and D. Guitton, personal communication; W.H.R., J. White and D.L.S., unpublished observations), inactivation of a subset of the active population produced reductions in saccadic velocity.

The population-averaging hypothesis holds that computation of the metrics of saccadic eye movements is based on a weighted average of the saccade-related discharges of the entire active population. According to the hypothesis, small changes in the direction and/or amplitude of saccadic movements are produced by slight shifts in the locus of the active population in the superior colliculus. The effects of variability or 'noise' in the discharge frequency of a particular neuron are minimized as the contribution of each neuron to the direction and amplitude of the movement is relatively small. In this manner, broadly tuned movement fields (resulting in a large population of neurons being active during a specific movement) may contribute to, rather than detract from saccadic accuracy.

We acknowledge invaluable programming assistance from Kathy Pearson and technical help from Isabel Ragland. This work was supported by the National Eye Institute of the National Institutes of Health.

Received 11 December 1987; accepted 3 February 1988.

1. Sparks, D. L. *Physiol. Rev.* **66**, 118-171 (1986).
2. Sparks, D. L., Holland, R. & Guthrie, B. L. *Brain Res.* **113**, 21-34 (1975).
3. McIlwain, J. T. *Int. Rev. Physiol.* **10**, 223-248 (1976).
4. Van Gisbergen, J. A. M., Van Opstal, A. J. & Tax, A. A. M. *Neuroscience* **21**, 541-555 (1987).
5. Deubel, H., Wolf, W. & Hauske, G. in *Theoretical and Applied Aspects of Eye Movement Research* (eds Gale, A. G. & Johnson, F.) 55-62 (Elsevier, Amsterdam, 1984).
6. Grossberg, S. & Kuperstein, M. *Neural Dynamics of Adaptive Sensory-Motor Control: Ballistic Eye Movements* (North-Holland, Amsterdam, 1986).
7. Wurtz, R. H. & Goldberg, M. E. *J. Neurophysiol.* **35**, 575-586 (1972).
8. Schiller, P. H. & Stryker, M. J. *Neurophysiol.* **35**, 915-924 (1972).
9. Robinson, D. A. *Vision Res.* **12**, 1795-1808 (1972).
10. Sparks, D. L. *Brain Res.* **156**, 1-16 (1978).
11. Hikosaka, O. & Wurtz, R. H. *J. Neurophysiol.* **53**, 266-291 (1985).
12. Hikosaka, O. & Wurtz, R. H. *Expl. Brain Res.* **61**, 531-539 (1986).
13. Berthoz, A., Grantly, A. & Droulez, J. *Neurosci. Lett.* **72**, 289-294 (1986).
14. Malpeli, J. G. & Schiller, P. H. *J. Neurosci. Meth.* **1**, 143-151 (1979).
15. Fuchs, A. F. & Robinson, D. A. *J. Appl. Physiol.* **21**, 1068-1070 (1966).

Basic fibroblast growth factor prevents death of lesioned cholinergic neurons *in vivo*

Kevin J. Anderson, David Dam, Sandra Lee & Carl W. Cotman*

Department of Psychobiology, University of California, Irvine, California 92717, USA

Cutting the axons of the cholinergic neurons that project to the hippocampal formation results in death of most of these cells^{1,2}. Previous studies have shown that administration of nerve growth factor before or at the same time as the lesion will prevent this cell death^{3,4}. Here we demonstrate that basic fibroblast growth factor (FGF) administered into the brain reduces the death of cholinergic neurons in the medial septum and diagonal band of Broca after transection of their axons, in both young adult and aged rats. Moreover, FGF can partially protect against death of cholinergic neurons even when administered two days after axonal transection. These results indicate a possible function for FGF in the normal support of basal forebrain cholinergic neurons, but its range of activity could be wider, for FGF also supports non-cholinergic neurons *in vitro*^{5,6}, it is localized in many of the central nervous system neurons⁷, and it is found in relatively high concentrations in the brain⁸.

Fimbria-fornix transection axotomizes the cholinergic neurons that project to the hippocampal formation, and results in the death of up to one-half of the cholinergic neurons in the

medial septum and vertical limb of the diagonal band of Broca in untreated animals². Twelve days following the fimbria-fornix transection, the animals were anaesthetized, killed by intracardiac perfusion, and processed for acetylcholinesterase (AChE) histochemistry and cresyl violet staining. There is a close correlation between AChE-positive neurons and those which contain choline acetyltransferase in the septum, so AChE staining is a reliable marker for cholinergic neurons in this brain area⁹. All AChE-positive neurons in the medial septum and diagonal band were counted in ten sections from each subject. The number of cholinergic neurons were plotted as a percentage of the control (contralateral) septum.

Examination of the medial septum and diagonal band region twelve days after fimbria-fornix transection showed a dramatic loss in the large AChE-positive neurons ipsilateral to the lesion (Fig. 1). Fimbria-fornix transection produced a loss of 58% of the cholinergic neurons in the medial septum and 46% of cholinergic cells in the diagonal band (Fig. 2). There was no loss of neurons in the horizontal limb of the diagonal band or the substantia innominata. Examination of alternate cresyl violet-stained sections also showed an obvious reduction in the number of stained neurons (42% decrease in the medial septum and 25% in the vertical limb of the diagonal band), indicating that the loss of cholinergic neurons cannot be accounted for by a decrease of AChE within individual cells. These values compare to those obtained in previous studies²⁻⁴.

In FGF-treated subjects, there was an increase in the number of large AChE-positive neurons relative to the lesioned side in controls. No obvious morphological differences were noted between the neurons on the lesioned side when compared to the cells in the contralateral septum. In sections from young adult animals that had received FGF infusions, 80% of the medial septum and 88.4% of the diagonal band cholinergic neurons survived the fimbria-fornix transection (Figs 1 and 2). Thus, FGF rescued 66% of the medial septum and 74% of the diagonal band cholinergic neurons which otherwise would have died ($P < 0.025$). No significant differences were seen between the unlesioned side of FGF-treated subjects when compared to the unlesioned side of control subjects.

In aged subjects (24-26-month-old rats), FGF infusion rescued 45% of the diagonal band and 47% of the medial septal neurons that would have died ($P < 0.025$). When FGF was infused two days after the fimbria-fornix transection in young adult animals, a higher percentage of cholinergic neurons survived axotomy. In these cases, a higher dose ($5.0 \mu\text{g ml}^{-1}$) was effective, whereas a lower dose ($0.5 \mu\text{g ml}^{-1}$) was ineffective. FGF facilitated the survival of cholinergic neurons in the vertical limb of the diagonal band (74% rescued, $P < 0.025$), but it did not significantly improve survival in the medial septum. It is clear then that FGF can be used after injury to rescue dying neurons, but the minimal time for complete protection and the maximal time for FGF to be still effective after injury have not been established.

FGF therefore promotes survival of brain neurons subjected to axonal transection, and also of axotomized retinal ganglion cells¹⁰, indicating that this molecule can have pronounced neurotrophic effects on central neurons *in vivo*. FGF is the second neurotrophic molecule which elicits neuron survival when administered exogenously and it is effective at extremely low doses, $< 6 \text{ ng}$ having been injected daily in the pre-treated group. These findings significantly extend the activity of FGF in the brain. The localization of FGF to neurons of the hippocampal formation⁷, the primary target of the cholinergic neurons of the medial septum and diagonal band, suggests that FGF may function in the normal maintenance or function of the septo-hippocampal pathway. Whether exogenously administered or endogenous FGF is acting directly on septal cholinergic neurons is not clear. FGF may act indirectly on glia^{11,12}, which in turn may provide trophic substances to the injured septal cholinergic neurons.

* To whom correspondence should be addressed.

Fig. 1 Photomicrographs of AChE stained neurons taken from *a*, a naïve control animal. *b*, A lesion control subject and *c*, a subject with a fimbria-fornix lesion that had received an infusion of FGF. *d-f*, Cresyl violet-stained sections of the medial septal nucleus from the same groups as above. All are shown 12 d after unilateral fimbria-fornix transection and all lesion subjects are shown with the lesion side on the right. Arrows indicate midline of section. Note the prominent unilateral loss of neurons in the lesion subjects. Neuronal death was reduced in the FGF-treated animals.

Methods. Rats (Sprague-Dawley, 3-4 months of age) received infusions of FGF (AMGen, California). This product was prepared by a recombinant DNA procedure and purified by sequential chromatography. Specific activity of FGF was 20-50 pg ml⁻¹ protein for ED₅₀ in 3T3 fibroblasts. FGF was diluted to 500 ng ml⁻¹ in sterile saline containing 0.1% v/v autologous rat serum as carrier. Controls received sterile saline with 0.1% autologous rat serum (*n* = 4), or no infusion (*n* = 6). Values for these two groups were within 5% and there were no statistical differences between the groups (*P* > 0.1). Thus they were combined for statistical comparisons with treated subjects (*P* > 0.45). Unlesioned control subjects (*n* = 4) were also examined. All subjects were anaesthetized with Chlorpент (50 mg kg⁻¹; Abbott Labs) before surgical procedures. Infusions were carried out by placing a cannula (25g, stainless steel) into the right lateral ventricle and attaching an osmotic mini-pump (model 2002, Alza, Palo Alto) with a flow rate of 0.5 µl h⁻¹ for 14 d. Two days after cannula implantation, the animals were subjected to a unilateral (right) transection of the fimbria-fornix and supracollosal striae, as previously described¹⁴. A separate group of animals were implanted with mini-pumps containing FGF (500 ng ml⁻¹, *n* = 4; or 5.0 µg ml⁻¹, *n* = 3) 2 d after fimbria-fornix transection. Animals were allowed to survive for 12 d, anaesthetized and processed for acetylcholinesterase histochemistry¹⁵. Animals were treated with diisopropylfluorophosphate (Sigma, 2 mg kg⁻¹ in sterile saline) 1 h before killing. Sections were stained for AChE using tetraisopropyl-pyrophosphoramidate (100 µM) as an inhibitor of non-specific esterases.

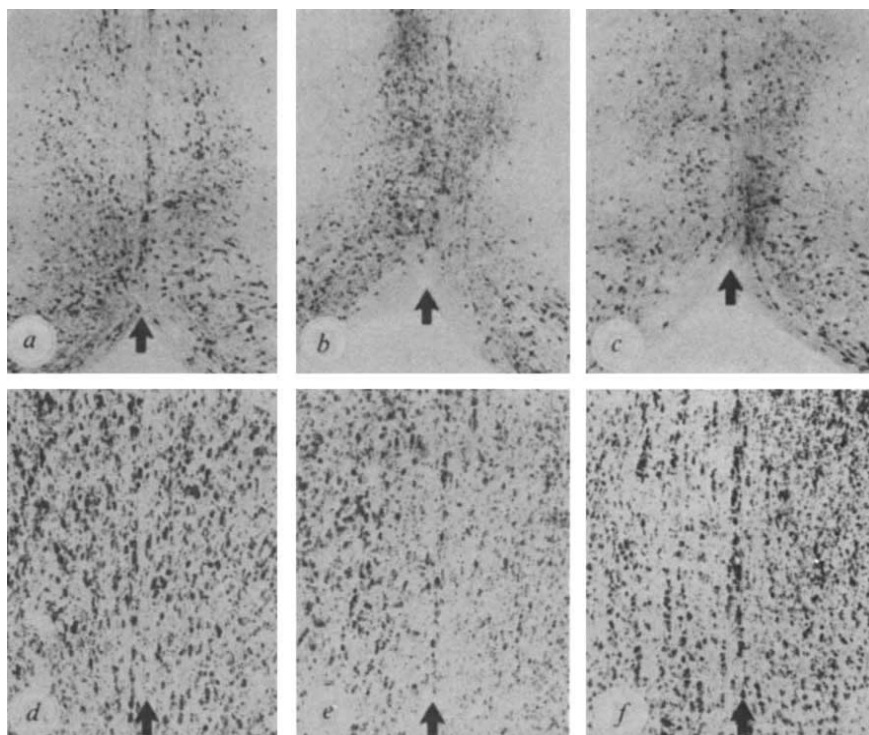
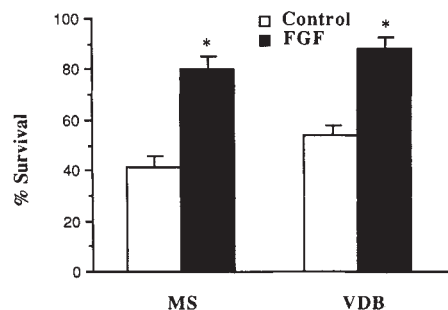


Fig. 2. Percent of cholinergic neurons in lesioned relative to unlesioned medial septum (MS) and vertical limb of the diagonal band (VDB) in FGF-treated and untreated (control) animals. (**P* < 0.025). Sections (30 µm) were taken through the rostro-caudal extent of the septal nuclear area and through the lesion site to ensure the lesion was complete. Every fourth section was used for counting AChE-positive neurons. Boundaries for the regions of the septal nuclear area were as described². Cells were counted blind and analysed as percentage of the control (contralateral) septal nuclear area. A one-way analysis of variance was used to determine if there were differences between the groups, and Scheffé *F*-tests were used for post-hoc analyses.



Our evidence suggests that FGF is also effective in aged animals, where neurons are presumably more sensitive to injury. Chronic infusions of nerve growth factor (NGF) into aged rats will improve performance on a spatial memory task¹³. This behavioural improvement was attributed to effects of nerve growth factor on cholinergic neurons. It remains to be seen whether FGF can also ameliorate behavioural impairments in aged or young animals by its action on intact or axotomized cholinergic neurons.

As NGF and FGF apparently both exist in the hippocampal formation, these factors could act independently or together to affect the septo-hippocampal pathway. Because NGF promotes cholinergic neuron survival, it has been proposed as a possible therapy in neurodegenerative disorders such as Alzheimer's disease. FGF itself, or drugs acting on FGF receptors, either alone or in conjunction with NGF, might also prove to be beneficial in the treatment of Alzheimer's disease. FGF also acts on a wide variety of non-cholinergic neurons *in vitro*, including cortical and hippocampal neurons^{5,6}, and could possibly be useful in other neurodegenerative disorders of the CNS involving loss of non-cholinergic neurons, such as stroke, epilepsy or

ischaemia. The survival-promoting activity of FGF may also be applicable in the treatment of neural trauma resulting from spinal cord or head injury.

Received 31 December 1987; accepted 2 February 1988.

1. Daitz, H. M. & Powell, T. P. S. *J. Neurol. Neurosurg. Psychiatr.* **17**, 75-82 (1954).
2. Gage, F. H. *et al. Neuroscience* **19**, 241-256 (1986).
3. Hefti, F. *J. Neurosci.* **6**, 2155-2162 (1986).
4. Williams, L. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9231-9235 (1986).
5. Walicke, P., Cowan, W. M., Ueno, N., Baird, A. & Guillemain, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3012-3016 (1986).
6. Morrison, R. S., Sharma, A., de Vellis, J. & Bradshaw, R. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7537-7541 (1986).
7. Pettmann, B., Labourdette, G., Weibel, M. & Sensenbrenner, M. *Neurosci. Lett.* **68**, 175-180 (1986).
8. Gospodarowicz, D., Cheng, J., Lui, G.-M., Baird, A. & Böhlent, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6963-6967 (1984).
9. Mesulam, M.-M., Mufson, E. J., Wainer, B. H. & Levey, A. I. *Neuroscience* **10**, 1185-1201 (1983).
10. Sievers, J., Hausmann, B., Unsicker, K. & Berry, M. *Neurosci. Lett.* **76**, 157-162 (1987).
11. Pruss, R. M., Bartlett, P. F., Gavrilovic, J., Lisak, R. P. & Rattray, S. *Dev. Brain Res.* **2**, 19-35 (1982).
12. Gospodarowicz, D., Neufeld, G. & Schweigerer, L. *Cell Differ.* **19**, 1-17 (1986).
13. Fischer, W. *et al. Nature* **329**, 65-68 (1987).
14. Anderson, K. J., Gibbs, R. B., Salvaterra, P. M. & Cotman, C. W. *J. Comp. Neurol.* **249**, 279-292 (1986).
15. Butcher, L. L. in *Handbook of Chemical Neuroanatomy* (eds Björklund, A. & Hokfelt, T.) 1-49 (Elsevier, Amsterdam, 1983).

Stimulus-dependent myristoylation of a major substrate for protein kinase C

Alan A. Aderem, Katherine A. Albert, Matthew M. Keum,
James K. T. Wang, Paul Greengard & Zanvil A. Cohn

The Rockefeller University, 1230 York Avenue, New York,
New York 10021, USA

Bacterial lipopolysaccharide (LPS), the major surface component of gram-negative bacteria, exerts a profound effect on the immune system by enhancing the release of proteins and arachidonic acid metabolites from macrophages (for review see ref. 1). The molecular mechanism(s) by which LPS induces these various secretory responses is unknown. We previously reported that LPS promotes the myristoylation of several macrophage proteins including one with a relative molecular mass (M_r) of 68K². We have now found that by several criteria the 68K myristoylated protein is similar or identical to the 80/87K protein, a major specific substrate for protein kinase C (PKC) found in brain and fibroblasts³⁻⁶ (for review see refs 7, 8). We have also found that the myristoylated PKC substrate is quantitatively associated with the membrane fraction. Myristoylation of the PKC substrate may target it to the membrane and constitute a transduction pathway for stimulus-response coupling.

When murine resident peritoneal macrophages are treated with LPS, ³H-myristic acid is specifically incorporated into several proteins, one of which has an apparent molecular mass of 68K (ref. 2 and Fig. 1, lysate). This 68K protein can be immunoprecipitated with an antibody developed against the 87K protein, a specific PKC substrate purified from bovine brain^{4,9}, but not with pre-immune serum (Fig. 1).

Whereas LPS stimulates the incorporation of ³H-myristic acid into the 68K protein, TPA (12-*o*-tetradecanoyl phorbol-13-acetate) does not (Fig. 1). TPA does, however, promote the phosphorylation of the 68K protein in intact macrophages (Fig. 2a), supporting the contention that it is a PKC substrate. LPS also promotes the phosphorylation of the 68K protein, albeit to a lesser extent than TPA (Fig. 2a). LPS-dependent phosphorylation of the 68K protein occurs more slowly than does TPA-dependent phosphorylation (half-maximal phosphorylation with LPS at 60 min, half-maximal phosphorylation with TPA at 7 min; data not shown). Furthermore, the maximal level of phosphorylation obtained with LPS was approximately 50% of that obtained with TPA.

The apparent molecular mass of the brain and fibroblast 80/87K protein has been reported to vary with the species and with the analytical gel system used^{4,5,7,8}. The mouse macrophage, mouse brain, rat brain, and bovine brain proteins were electrophoresed together in two gel systems. On a 6–12% linear gradient SDS polyacrylamide gel (Nevilles buffer, ref. 2), the apparent M_r values obtained were 68K, 69K, 70K, and 76K, respectively (Fig. 2c). On an 8% SDS polyacrylamide gel (Laemmli buffer, ref. 4), the M_r values were 79K, 80K, 83K, and 87K, respectively (data not shown). The variability in apparent M_r observed with different gel systems may be explained by hydrodynamic studies of the bovine brain 87K protein which yield a calculated molecular mass of 68 kDa and suggest that the protein is a very asymmetric, highly elongated monomer⁹. When subjected to isoelectric focussing, the phosphorylated and myristoylated protein from murine macrophages and the phosphorylated proteins from murine, rat and bovine brain were found to be acidic with a pI of approximately 4.5 (ref. 4, data not shown). Furthermore, the 68K murine macrophage protein, phosphorylated intracellularly in response to TPA, and the purified 87K bovine brain protein, phosphorylated *in vitro* with purified PKC, generated identical phos-

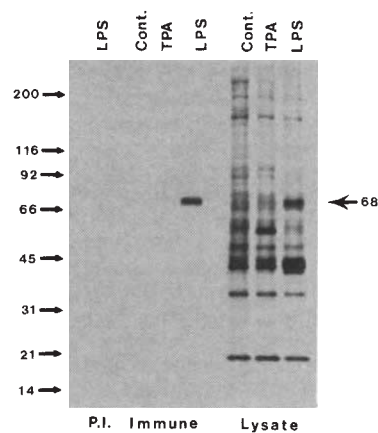


Fig. 1 Fluorogram showing that LPS induces the myristoylation of a 68K PKC substrate in macrophages. Murine resident macrophages were isolated by peritoneal lavage and treated with ³H myristic acid in the absence (control) or presence of 10 ng ml⁻¹ *S. Abortus Equi* LPS or 10 ng ml⁻¹ TPA. After 2 h the cells were lysed in a NP40 based solution containing protease inhibitors (phenylmethylsulfonyl fluoride, 1 mM; aprotinin, 0.28 trypsin inhibitor units per ml; diisopropylfluorophosphate, 1 mM; EDTA, 15 mM) and phosphatase inhibitors (potassium fluoride, 50 mM; sodium phosphate, 50 mM; sodium pyrophosphate, 10 mM). Myristoylated proteins were analysed in the lysate and on fractions which had been immunoprecipitated using an antiserum directed against the bovine 87K protein or preimmune serum. Myristoylated proteins were visualized by fluorography after SDS-PAGE on 6–12% gels.

Methods. All methods were as described previously^{2,4,9} except where noted in figure legends. For preparation of the antibody^{4,9}, 75 µg of the purified bovine brain 87K protein (a single band on a silverstained gel) was emulsified in complete Freund's adjuvant and injected into multiple sites in a New Zealand female rabbit. The rabbit was given a booster injection with 50 µg of the protein at 7 weeks after the first injection and bled at weekly intervals thereafter. The bleeding at 12 weeks was used for the present study. The antiserum was specific for the 87K PKC substrate⁴. The studies described here were carried out with the IgG fraction of the serum. Figs 1 and 2 were identical when repeated with affinity purified IgG.

phopeptide maps when subjected to limited proteolysis with *S. aureus* V8 protease (Fig. 2b). Both proteins generated major and minor phosphopeptides of 13K and 9K, respectively. The two-dimensional phosphopeptide maps following thermolytic digestion of the immunoprecipitated murine macrophage and purified bovine brain proteins were also identical (data not shown). Thus a variety of data suggest that the 68K PKC substrate in macrophages and the 80/87K PKC substrate in brain are closely related structurally.

To demonstrate that myristic acid is the actual lipid species linked to the 68K protein, we released the fatty acid with methanolic HCl and then analysed it by high pressure liquid chromatography. Virtually all of the radiolabel (>98%) co-eluted with myristic acid or methyl myristic acid (Fig. 3). The myristic acid-protein linkage was resistant to 1M hydroxylamine (pH 10), suggesting that the fatty acid moiety is bound to the protein by an amide bond (data not shown).

We have compared the distribution between the membrane and cytosolic fractions of the myristoylated and phosphorylated 68K PKC substrate. The major portion (>90%) of the myristoylated 68K protein was found in the membrane fraction, whereas the major portion (>75%) of the phosphorylated 68K protein was observed in the cytosolic fraction (Fig. 4). As one working hypothesis, we propose the following sequence of events to explain the observed distributions of the myristoylated and phosphorylated forms of the 68K protein. (1) LPS promotes the

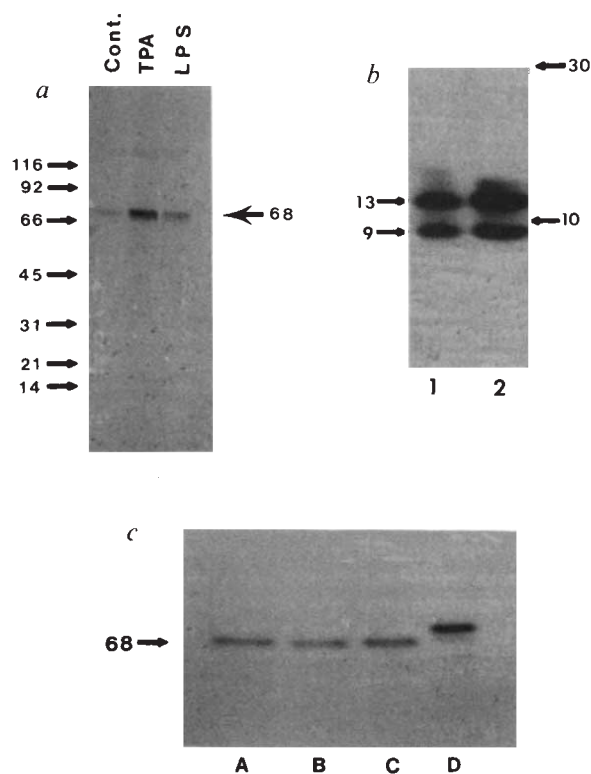


Fig. 2 *a*, Autoradiogram showing that TPA and LPS promote the phosphorylation of the 68K PKC substrate in macrophages. Murine resident peritoneal macrophages were preincubated in phosphate-free minimal Eagles medium containing 250 μCi $^{32}\text{P}_i$ for 30 min followed by incubation in the absence (control) or presence of TPA (10 ng ml^{-1}) or LPS (10 ng ml^{-1}). After 30 min the cells were scraped into lysis buffer containing protease and phosphatase inhibitors and immunoprecipitated with antiserum directed against the bovine 87K protein^{4,9}. The immunoprecipitated phosphoprotein was visualized by autoradiography following SDS-PAGE on a 6–12% gel. *b*, Autoradiogram showing phosphopeptide maps after limited proteolysis with *S. aureus* V8 protease. Lane 1, purified bovine brain 87K protein phosphorylated *in vitro* with purified PKC (refs 4, 9). Lane 2, murine macrophage 68K protein phosphorylated intracellularly by addition of TPA. Both phosphorylated proteins were immunoprecipitated with the bovine 87K protein antiserum and subjected to SDS-PAGE as described above. The phosphorylated protein, located by autoradiography, was excised and subjected to limited proteolysis during SDS-PAGE on 15% gels^{4,25}. The 10K and 30K molecular weight markers were the lower and upper site phosphopeptide markers obtained by digestion of synapsin I by the same enzyme²⁵. *c*, Autoradiogram showing molecular mass heterogeneity of phosphorylated PKC substrates immunoprecipitated with bovine 87K protein antiserum from murine macrophage (lane A), murine brain (lane B), rat brain (lane C) and bovine brain (lane D). Murine macrophage proteins were phosphorylated by treating the cells with TPA. Post-mitochondrial supernatants of murine and rat brain homogenates and purified bovine brain 87K protein were phosphorylated *in vitro* with purified PKC as described previously^{4,9}. The phosphoproteins were immunoprecipitated with an antiserum directed against the bovine protein and visualized by autoradiography following SDS-PAGE on a 6–12% gel.

myristoylation of the 68K protein. (2) The myristoylated protein moves to the membrane where it is more closely associated with PKC (known to be active at the membrane¹⁰). (3) PKC phosphorylates the 68K protein. (4) The phosphorylated 68K protein is released to the cytoplasm (possibly as a result of demyristoylation).

This hypothetical sequence of events is supported by several lines of evidence. First, in all systems described thus far myris-

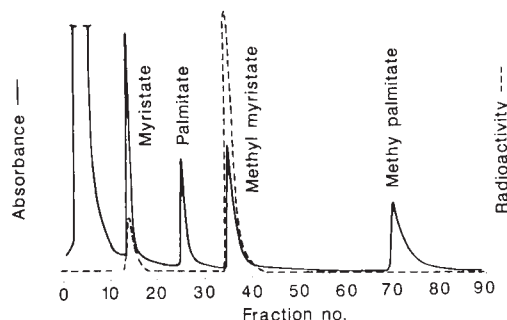


Fig. 3 Analysis of protein-bound fatty acid after acid methanolysis. Macrophages were incubated with ^3H -myristate and LPS for 2 h after which the ^3H -labelled 68K protein was immunoprecipitated as described in the legend to Fig. 1. Lipids were removed from the protein²⁶, subjected to acid methanolysis, and the fatty acid methyl esters were isolated by reverse phase HPLC (ref. 2). All the radioactivity co-eluted with the myristate and methyl myristate standards. The positions on the chromatogram of the fatty acid standards are indicated by the absorbance peaks. The methyl myristate absorbance peak corresponds to an absorbance of 0.8 o.d. and the radioactivity associated with this peak corresponds to 7674 d.p.m.

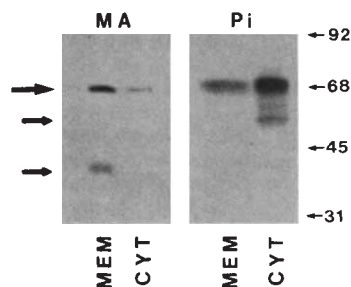


Fig. 4 Subcellular distribution of myristoylated and phosphorylated 68K protein. Macrophages were incubated with ^3H -myristate and LPS (MA) or with $^{32}\text{P}_i$ and LPS (Pi) for 90 min. TPA was then added for an additional 30 min. The concentrations and specific activities of the reagents were as described in the legends to Figs 1 and 2. The cells were then separated into membrane (MEM) and cytosol (CYT) fractions (ref. 2). Immunoprecipitated protein was visualized by autoradiography following SDS-PAGE on a 6–12% gel.

toylation appears to occur during or very soon after translation (for review see ref. 27), and this also appears to be the case in macrophages, since LPS-induced myristoylation of the 68K protein is rapidly blocked by protein synthesis inhibitors (data not shown). Second, in at least two instances protein-bound myristic acid has been shown to have a role in targeting proteins to the membrane. Myristoylation of p60^{src} and Moloney murine leukaemia virus p65^{gag} is necessary for plasma membrane association, transforming potential and viral budding^{11–14}. Third, pretreatment of macrophages with LPS results in greatly increased TPA-dependent phosphorylation of the 68K protein and also results in a shift in the TPA dose-response curve such that half-maximal phosphorylation of p68 in LPS-treated macrophages is obtained with 10-fold lower concentrations of TPA than those required in macrophages not treated with LPS (A. Rosen and A. Aderem; manuscript in preparation). Fourth, experiments with ^3H -myristic acid and ^{35}S -methionine indicate a more rapid turnover in response to TPA of myristic acid than of protein, raising the possibility that TPA may induce the demyristoylation of the 68K protein (A. Rosen and A. Aderem, preliminary results). An alternative, and in our view less likely,

explanation for the differential distribution of the myristoylated and phosphorylated p68 is that a non-myristoylated species of p68 is phosphorylated in the cytoplasm by a fragment of PKC which is released from the membrane by proteolysis¹⁰.

It is interesting to consider the possible biological significance of this dual regulatory system which results in the 68K protein being modified with myristic acid and phosphate. A clue to its significance comes from our observation that LPS primes macrophages for a greatly enhanced response (increased arachidonic acid metabolism) to subsequent stimulation by TPA (ref. 15). Coincident with the establishment of the primed state, LPS promotes the myristoylation of three proteins, including the 68K PKC substrate². We speculate that the LPS-induced, myristic acid-dependent targeting of the 68K PKC substrate to the membrane facilitates the transduction of the subsequent, PKC-dependent stimulus, resulting in the potentiated secretion of arachidonic acid metabolites.

The function of this major PKC substrate is not yet known. It was first described as an 87K protein in rat brain synaptosomes where it was found to be phosphorylated in response to phorbol esters and potassium depolarization^{3,16,17}. Our preliminary data indicate that this protein is also myristoylated in murine frontal cortex cells. The 87K PKC substrate and the 68K macrophage protein are similar or identical to the 80K fibroblast protein whose phosphorylation is promoted by phorbol esters and by growth factors such as platelet derived growth factor, fibroblast growth factor, serum and bombesin^{6,18-24}. It would be of interest to know whether myristoylation, like phosphorylation, is involved in all of these signal transduction systems. The convergence of these two signal transduction pathways via the post-translational modification of a PKC substrate with myristic acid would represent a novel mechanism for the potentiation of PKC-dependent signals.

This work was partially supported by National Institutes of Health Grant AI 25032 and MH 3937 and by a grant-in-aid from the Squibb Institute for Medical Research. A.A.A. is a Pew Scholar in the Biomedical Sciences.

Received 25 August 1987; accepted 20 January 1988.

- Morrison, D. C. & Ulevitch, R. J. *Am. J. Path.* **93**, 527-617 (1978).
- Aderem, A. A., Keum, M. M., Pure, E. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5817-5821 (1986).
- Wu, W. S., Walaas, S. I., Nairn, A. C. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5249-5253 (1982).
- Albert, K. A., Walaas, S. I., Wang, J. K.-T. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2822-2826 (1986).
- Blackshear, P. J., Wen, L., Glynn, B. P. & Witters, L. A. *J. bio. Chem.* **261**, 1459-1469 (1986).
- Rozengurt, E., Rodriguez-Pena, M. & Smith, K. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7244-7248 (1983).
- Neidel, J. E. & Blackshear, P. J. in *Phosphoinositides and receptor mechanisms* (ed. Putney, J. W.) Ch. 3, 47-88 (Alan R. Liss, New York, 1986).
- Woodgett, J. R., Hunter, T. & Gould, K. L. in *Cell Membranes: Methods and Reviews*, 3rd edn (eds Elson, E. L., Frazier, W. A. & Glazer, L.) (Ch. 6, 215-340, Plenum, New York, 1987).
- Albert, K. A., Nairn, A. C. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7046-7050 (1987).
- Nishizuka, Y. *Science* **233**, 305-312 (1986).
- Pellman, D., Garber, E. A., Cross, F. R. & Hanafusa, H. *Nature* **314**, 374-377 (1985).
- Cross, F. R., Garber, E. A., Pellman, D. & Hanafusa, H. *Molec. cell Biol.* **4**, 1834-1842 (1984).
- Kamps, M. P., Buss, J. E. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4625-4628 (1985).
- Rein, A., McClure, M. R., Rice, N. R., Luftig, R. B. & Schultz, A. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7246-7250 (1986).
- Aderem, A. A., Cohen, D. S., Wright, S. D. & Cohn, Z. A. *J. Exp. Med.* **164**, 165-179 (1986).
- Wang, J. K.-T., Walaas, S. I. & Greengard, P. *J. Neurosci.* **8**, 281-288 (1988).
- Nichols, R. A., Haycock, J. W., Wang, J. K.-T. & Greengard, P. *J. Neurochem.* **48**, 615-621 (1987).
- Blackshear, P. J., Witters, L. A., Girard, P. R., Kuo, J. F. & Quamo, S. N. *J. bio. Chem.* **260**, 13304-13315 (1985).
- Rodriguez-Pena, A. & Rozengurt, E. *EMBO J.* **4**, 71-76 (1985).
- Tsuda, T., Kaibuchi, K., Kawahara, Y., Fukuzaki, H. & Takai, Y. *FEBS Lett.* **191**, 205-210 (1985).
- Isacke, C. M., Meisenhelder, J., Brown, K. D., Gould, K. L., Gould, S. J. & Hunter, T. *EMBO J.* **5**, 2889-2898 (1986).
- Sato, C., Nishizawa, K., Nakayama, T. & Kobayashi, T. *J. Cell Biol.* **100**, 748-753 (1985).
- Coughlin, S. R., Lee, W. M. F., Williams, P. W., Giers, G. M. & Williams, L. T. *Cell* **43**, 243-251 (1985).
- Mahadevan, L. C., Aitken, A., Heath, J. & Foulkes, J. G. *Embo J.* **6**, 921-926 (1987).
- Huttner, W. B. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5402-5406 (1979).
- Olson, F. N., Towler, D. A. & Glaser, L. *J. bio. Chem.* **260**, 3784-3790 (1985).
- Sefton, B. M. & Buss, J. E. *J. Cell Biol.* **104**, 1449-1453 (1987).

Intracellular free calcium rise triggers nuclear envelope breakdown in the sea urchin embryo

Richard A. Steinhardt & Janet Alderton

Department of Zoology, University of California, Berkeley, California 94720, USA

Cytosolic free calcium has recently been implicated in the regulation of mitosis in plant and animal cells¹⁻⁸. We have previously found correlations between increases in the levels of intracellular free calcium $[Ca^{2+}]_i$ and visible transitions of structure at nuclear envelope breakdown (NEBD) and the onset of anaphase during mitosis in sea urchin embryos and tissue culture cells^{7,8}. To go beyond correlations it is necessary to manipulate $[Ca^{2+}]_i$, and in sea urchin embryos this requires the injection of calcium-chelator buffer solutions as the changes in free calcium in the cell cycle are dependent on intracellular stores^{7,9}. We report here that blocking the increase in $[Ca^{2+}]_i$ which just precedes NEBD prevents this from taking place and halts mitosis. Subsequent injections which momentarily increase $[Ca^{2+}]_i$, or a natural recovery of the higher calcium levels, result in NEBD and the successful continuation of mitosis. Similarly, artificially increasing calcium by early injections results in early NEBD. We conclude that the increase in $[Ca^{2+}]_i$ preceding NEBD is an essential regulatory step required for entry into mitosis.

Previously we calculated the mean post-fertilization baseline of $[Ca^{2+}]_i$ measured with the fluorescent tetracarboxylate chelator fura-2 to be 277 ± 28 nM and the elevation of $[Ca^{2+}]_i$ just preceding NEBD was shown to reach an average peak value of 404 ± 20 nM⁷. Here, also using the fura-2 method, the baseline $[Ca^{2+}]_i$ measured after activation by fertilization or ammonia treatment averaged 138 ± 17 nM ($n=9$) and the peak associated with NEBD averaged 566 ± 108 nM ($n=9$).

The injection of calcium-chelator buffer solutions along with the indicator fura-2 prevents the transient increase in $[Ca^{2+}]_i$ immediately preceding NEBD and blocks both NEBD and further progress of mitosis. Figure 1 illustrates these correlations showing that blocking the rise in free calcium stops the progress of mitosis following ammonia activation or fertilization in *Lytechinus pictus* embryos. In Fig. 1a the normal rise in $[Ca^{2+}]_i$ precedes NEBD. In both fertilization and ammonia activation, NEBD was observed within 1-3 min following the first calcium rise occurring more than 40 min post-activation. In Fig. 1b $[Ca^{2+}]_i$ is buffered with EGTA preventing the rise and blocking NEBD. Figure 1c follows $[Ca^{2+}]_i$ after fertilization in another embryo which was watched continuously from the time of injection of the blocking buffer until $[Ca^{2+}]_i$ eventually recovered to the level which triggers NEBD. Table 1 summarizes 13 experiments in which we injected buffer solutions to test the hypothesis that the calcium peak is essential for NEBD. Keeping $[Ca^{2+}]_i$ at or below 270 nM with EGTA prevented NEBD. To control for possible toxic effects of organic acids we also used another chelator, nitrolotri-acetic acid (NTA) which is relatively ineffective at binding calcium¹⁰. This was found to have no effect on NEBD. To control for possible effects of pH we injected bis(aminophenoxy)ethane *N,N'*-tetraacetic acid (BAPTA)¹¹ instead of EGTA. We found this to be as effective at blocking NEBD as EGTA. As EGTA or BAPTA might block by binding a heavy metal such as zinc, we applied 10 μ M *N,N,N',N'*-tetrakis(2-pyridylmethyl)ethylenediamine (TPEN)¹² in a separate test and found that NEBD proceeded normally in 30 embryos observed continuously. Low concentrations of EGTA or BAPTA were ineffective (see Table 1). This threshold of ~ 0.2 - 0.3 mM of chelator also implies that the block is induced by binding calcium, not a heavy metal.

Further evidence that the block is due to calcium buffering

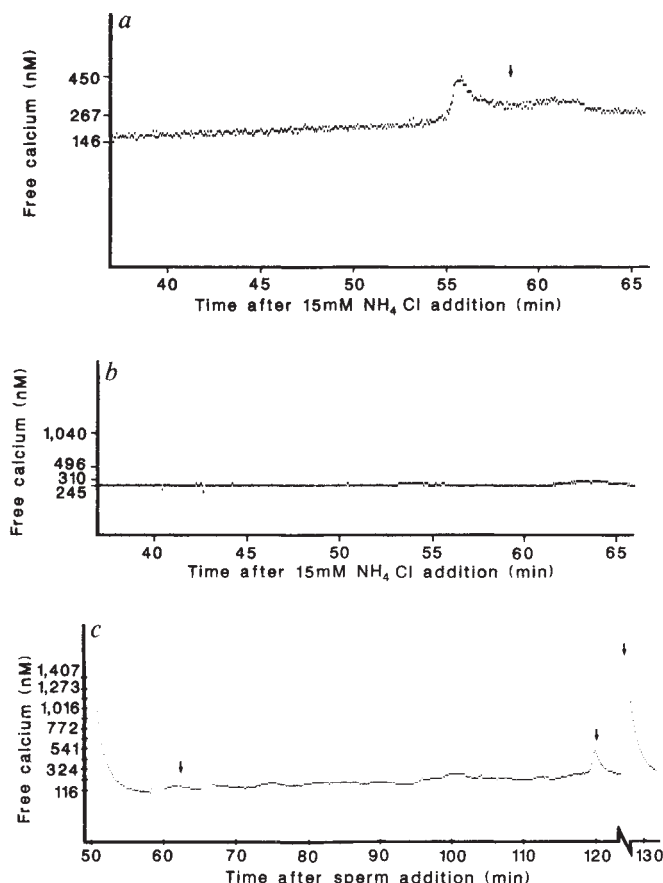


Fig. 1 *a*, Free intracellular calcium levels following activation by a 15 min long pulse with 15 mM ammonium chloride, pH 9.0. A typical example of the transient rise in $[Ca^{2+}]_i$ at the time of NEBD in ammonia-activated eggs of the sea urchin *L. pictus*. The trace before this rise is flat or slightly depressed compared with the resting levels before activation by ammonia and therefore has a much better signal-to-noise ratio for determining the relationship to NEBD than the more complex and generally elevated levels which typically follow activation by sperm⁷. The exact relationship of the time of NEBD (arrow) to the calcium rise has been determined by continuous observation under illumination with a narrow-bandwidth filter at 660 nm. The time of NEBD is defined as the end point at which cytoplasm enters the space which was formerly the clear area bounded by the nuclear envelope. *b*, A typical example in which both the calcium rise and NEBD are blocked by the injection of blocking buffer (80 mM EGTA, 160 mM potassium gluconate, 48 mM CaCl₂, 3.4 mM MgCl₂, 8 mM HEPES, pH 7.4) at a final intracellular buffer concentration of ~0.2 mM. *c*, Blocking the calcium rise after fertilization with blocking buffer also delays NEBD. The first arrow marks the time of NEBD in control eggs in the same dish. The second arrow marks the time of NEBD in the experimental egg after the calcium levels recover to the threshold peak for NEBD. The third arrow marks the injection of rescue buffer (without dye) to show that the original fura-2 is still reporting cytoplasmic free calcium. Rescue buffer is defined in the legend to Table 2. **Methods.** Eggs of *L. pictus* were microinjected with a solution of 27 mM fura-2 (Molecular Probes, Eugene, Oregon), 66 mM potassium gluconate, 5 mM CaCl₂, 5 mM HEPES, pH 7.3 to final concentrations in the range 5–50 μ M. Free calcium was calculated from the fluorescence ratio (R) obtained by dividing the fluorescence (500–530 nm) values from excitation at 350 nm by those at 385 nm. This ratio changes in the same direction as $[Ca^{2+}]_i$ and can be used to calculate $[Ca^{2+}]_i$ from the equation $K(R - R_0)/(R_s - R) = [Ca]$, where R_0 is the ratio at $[Ca] = 0$ and R_s is the ratio at saturating concentrations of calcium¹⁴. K represents $K_d(F_0/F_s)$, where K_d is the effective dissociation constant for fura-2 in the appropriate conditions (previously determined to be 7.74×10^{-7} , ref. 1), F_0 is the fluorescence at 385 nm in zero Ca and F_s is the fluorescence at 385 nm saturating calcium solutions. In these experiments, R_0 and R_s were modified by an average viscosity correction factor of 0.85 (ref. 8). R_0 and R_s were obtained from droplets of fura-2 in reference solutions containing 50 μ M fura-2, 155 mM KCl, 25 mM NaCl 100 mM Morpholinopropane Sulfonic Acid (MOPS) neutralized with KOH to pH 7.02 and either 10 mM K₂H₂EGTA or 1 mM CaCl₂. Excitation light was provided by interference filters (bandwidth = 5 nm) at 350 nm and 385 nm rotated by a stepping motor at 0.5–4 Hz. The emitted light was collected through a 500–530 nm interference filter and photon-counted in synchrony with the stepping motor so that the signals from the two excitation wavelengths were stored in separate memories of an IBM AT computer. Experiments were conducted at 18 °C.

Table 1 Effect of buffer injections on nuclear envelope breakdown

Expt.	Chelator concentration (mM) in cell	Baseline $[Ca^{2+}]_i$ after injection (nM)	$[Ca^{2+}]_i$ at time of control NEBD (nM)	Delay in NEBD from controls (min)	$[Ca^{2+}]_i$ at time of NEBD experimental (nM)
1	0.1 EGTA	—	—	0	—
2	0.2 EGTA	156	251	55	540
3	0.2 EGTA	176	—	45	—
4	0.3 EGTA	128	261	47	622
5	0.3 EGTA	—	—	47	—
6	1.0 EGTA	116	160	58	541
7	0.2 BAPTA	123	—	0	—
8	0.6 BAPTA	144	166	31	—
9	0.7 BAPTA	132	144	69	—
10	1.2 BAPTA	—	—	65	—
11	0.5 NTA	135	506	0	506
12	1.2 NTA	215	593	0	593
13	1.5 NTA	—	—	0	—

Injection solution compositions before dilution in cell: EGTA buffer, 80 mM EGTA, 160 mM potassium gluconate, 48 mM CaCl₂, 3.4 mM MgCl₂, 8 mM HEPES, pH 7.4; BAPTA buffer, 80 mM BAPTA, 80 mM potassium gluconate, 48 mM CaCl₂, 3.4 mM MgCl₂, 8 mM HEPES, pH 7.4; NTA buffer, 80 mM NTA (nitrilotriacetic acid), 160 mM potassium gluconate, 8 mM HEPES, pH 7.4. Final concentration inside cell is estimated by measurement of total cell fluorescence increase at 500–530 nm (excitation 350 nm) resulting from the inclusion of 5.2 mM fura-2 in the original injection solutions.

follows from an additional observation on the nature of the block. Figure 1c and Table 1 show that the EGTA or BAPTA block is temporary, inducing an average delay of 52 ± 4 min ($n = 8$). In three earlier experiments which were not followed all the way to NEBD, the block lasted for the entire period of observation (42–60 min). In experiments 2, 4 and 6, $[Ca^{2+}]_i$ was monitored continuously until the block of NEBD was overcome at 47, 55 and 58 min respectively. The values for $[Ca^{2+}]_i$ increased gradually reaching 540 nM, 620 nM and 541 nM respectively at the time of NEBD. Apparently EGTA loses its effectiveness at buffering calcium in the embryo with time. It is therefore possible that it becomes loaded from calcium stores, or extracellular calcium entry. Alternatively EGTA may be sequestered or lost from the cytoplasm so that its effective concentration becomes too low to buffer calcium. A similar phenomenon might explain the ineffectiveness of EGTA injections on the timing of cell division in *Xenopus* embryos¹³. But if EGTA can become ineffective with time how can we be sure that there is enough active fura-2 to accurately report the free calcium concentration? To answer this question we carried out six additional experiments in which eggs were injected with fura-2 before activation and then reinjected with rescue buffer (without additional dye) preset to higher calcium concentrations 1–2 h later. In every case the cytoplasmic calcium, as reported by the fura-2 previously injected, reached the micromolar levels normally seen when the fura-2 is injected at the same time as the buffer. The last arrow in Fig. 1c indicates where reinjection of rescue buffer (without additional dye) allows the levels of free calcium to reach micromolar levels. We conclude that there was sufficient dye to accurately report the free calcium levels in the cytoplasm at the end of these experiments.

It should be possible to overcome the EGTA block by subsequent reinjections of a rescue buffer solution preset to higher calcium concentrations. In fact, the act of reinjection itself, which is done in normal calcium-containing sea water to permit healing, instantly takes $[Ca^{2+}]_i$ to high levels. Normally NEBD is observed visually within 3 min after the transient $[Ca^{2+}]_i$ peak preceding it. The interval is slightly longer after rescue by

Table 2 Effect of blocking buffer injections followed by rescue buffer injections on nuclear envelope breakdown

First injection of blocking buffer (mM)	[Ca ²⁺] _i baseline after first injection (nM)	Total chelator after injection of rescue buffer (mM)	[Ca ²⁺] _i peak during rescue buffer injection (nM)	Interval to NEBD after rescue and total delay versus controls (min)
0.2	144	0.7	8065	12–27
0.3	161	0.5	2146	10–20
0.3	190	0.5	2028	7–13
0.4	—	0.6	—	11–22
0.5	118	1.0	4107	39–47

Blocking buffer composition equals the EGTA buffer (Table 1). Rescue buffer has 68 mM CaCl₂ substituted for the 48 mM CaCl₂ in the blocking buffer. The actual calcium peak is momentarily much higher at the site of the injection than that calculated from these concentrations due to calcium entry from the external solution of sea water during the injection procedure.

reinjection. This was successful in 4 of 5 experiments with NEBD occurring at an average of 10 ± 2.16 min after reinjection (Table 2). In one case, reinjection with rescue buffer failed to significantly shorten the period of the block. This anomaly may have resulted from an injury due to a poor injection.

We did 10 additional experiments to see if it was possible to advance the time of NEBD by injection with rescue buffer 46–71 min post-insemination to artificially induce an early calcium rise. In nine of these experiments NEBD followed within 3 min of the injection, advancing the time of NEBD by as much as 44 min for the earliest injection and advancing NEBD by the 12 minutes remaining to the time of the control NEBD for the latest injection done at 71 minutes post-insemination. In one case the injection had no effect with both control and experimental NEBDs occurring within 1 min of each other. Here, in contrast, when injections with blocking buffer were used, the calcium entry associated with the penetration of the injection pipette did not trigger an early NEBD in spite of the momentarily high levels of [Ca²⁺]_i. The only difference between blocking buffer and rescue buffer was the amount of added calcium. The calculated free calcium concentrations were 10^{-7} M and 4×10^{-7} M for the blocking buffer and the rescue buffer, respectively. As the egg was at a higher ionic strength than the solutions for which we know the binding constants for EGTA,¹⁰ we found free calcium concentrations to be slightly higher than predicted. The difference in free calcium between blocking buffer and rescue buffer is enough to trigger NEBD and closely matches the levels that correlate with normal NEBD in eggs which have not been injected with buffer.

We thank the NSF, the NIH and IBM for support, Dr Martin Poenie for discussions and Chester Regen for the work on the data acquisition program.

Received 24 November 1987; accepted 6 January 1988.

- Hepler, P. K. *J. cell Biol.* **100**, 1363–1368 (1985).
- Keith, C. H., Ratan, R., Maxfield, F. R., Baer, A. & Shelanski, M. L. *Nature* **316**, 848–850 (1985).
- Ratan, R., Shelanski, M. L. & Maxfield, F. R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5136–5140 (1986).
- Wolniak, S. K., Hepler, P. K. & Jackson, W. T. *J. cell Biol.* **96**, 598–605 (1983).
- Wolniak, S. M. & Bart, K. M. *Eur. J. cell Biol.* **39**, 273–277 (1986).
- Poenie, M. & Steinhardt, R. A. in *Calcium and Cell Function* Vol. 7 (ed. Cheung, W. Y.) 133–157 (Academic, New York, 1987).
- Poenie, M., Alderton, J., Tsien, R. Y. & Steinhardt, R. A. *Nature* **315**, 147–149 (1985).
- Poenie, M., Alderton, J., Steinhardt, R. & Tsien, R. *Science* **233**, 886–889 (1986).
- Chambers, E. L. *Eur. J. cell Biol.* **22**, 476 (1980).
- Martell, A. E. & Smith, R. M. *Critical Stability Constants* (Plenum, London, 1974).
- Tsien, R. Y. *Biochemistry* **19**, 2396–2404 (1980).
- Arslan, P., Di Virgilio, F., Beltrame, M. & Tsien, R. Y. *J. Biol. Chem.* **260**, 2719–2727 (1985).
- Baker, P. F. & Warner, A. E. *J. cell Biol.* **53**, 579–581 (1972).
- Gryniewicz, G., Poenie, M. & Tsien, R. Y. *J. Biol. Chem.* **260**, 3440–3450 (1985).

Translational control of InsP₃-induced chromatin condensation during the early cell cycles of sea urchin embryos

Jeremy Twigg, Rajnikant Patel & Michael Whitaker*

Department of Physiology, University College London, Gower Street, London WC1E 6BT, UK

* To whom correspondence should be addressed

The cycles of DNA synthesis and chromatin condensation in dividing cells are controlled by signals from the cytoplasm^{1–3}. Changes in the concentration of free calcium (Ca_i) in the cytoplasm control a variety of cellular functions⁴ and it has thus been suggested that observed variations in Ca_i during the cell cycle may be the cytoplasmic signal that co-ordinates nuclear and cytoplasmic division^{5–9}. We show here that increases in Ca_i induced by the calcium-releasing second messenger inositol 1,4,5-triphosphate (Ins(1,4,5)P₃), or by calcium buffers, cause premature chromatin condensation and breakdown of the nuclear envelope in sea urchin (*Lytechinus pictus*) early embryos. Both natural and induced chromatin condensation are prevented by calcium chelators. The nucleus becomes sensitive to the Ca_i signal 45 min after fertilization, but remains insensitive if protein synthesis is prevented. Our experiments demonstrate that Ca_i regulates the behaviour of the nucleus during the cell cycle, suggest that Ins(1,4,5)P₃ is a cell cycle messenger and indicate that there is an interaction between the protein and ionic signals that control the state of chromatin during the cell cycle.

A transient increase in Ca_i occurs at the time of nuclear envelope breakdown (NEB) and chromosome condensation in sea urchin early embryos⁷. We examined the influence of Ca_i on NEB and chromatin condensation by microinjecting calcium chelators and calcium buffers into newly fertilized sea urchin eggs and using a DNA-staining dye¹⁰ to monitor the state of the chromatin. Figure 1a shows that chromatin condensation, which normally occurs at 65 min after fertilization, can be prevented by prior microinjection of the calcium chelator

Table 1 Injection of calcium chelators and buffers, and Ins(1,4,5)P₃, into newly fertilized sea urchin eggs

	Cytoplasmic concentration	Chromatin condensation and NEB*	Time to cleavage
Agents blocking normal NEB and chromatin condensation†			
BAPTA	0.6 mM	18/19	100 min‡
BAPTA	1.4 mM	13/14	100–140 min‡
BAPTA	2.3 mM	0/18	failed to cleave
Injection controls	ND§	ND	100 min
Agents stimulating premature NEB and chromatin condensation			
CaEGTA	25 mM	5/5	190 min
(1 µM free Ca ²⁺)			
Injection controls	ND	0/8	180 min
Ins(1,4,5)P ₃	10 nM	0/4	ND
Ins(1,4,5)P ₃	30 nM	4/7	ND
Ins(1,4,5)P ₃	50 nM	24/27	100 min
InsP ₃	250 nM		
+		0/9	ND
EGTA	25 mM		

* Number of eggs showing either blockage of or premature NEB and chromatin condensation/number of eggs injected.

† Eggs were microinjected with 0.8 M BAPTA (0.8–3% egg volume).

‡ Half these eggs developed cleavage furrows but failed to cleave.

§ Not determined.

|| Second cell cycle microinjection and second cleavage.

Fig. 1 Chromosome condensation and NEB in eggs labelled with the vital DNA stain Hoechst 33342 (ref. 10) and microinjected with calcium buffers or $\text{Ins}(1,4,5)\text{P}_3$ at various times after fertilization. *a*, Left and middle: state of chromatin 75 min after fertilization in two embryos microinjected at 20 min with the calcium chelator BAPTA¹¹ to a final concentration of 2 mM. The example in the left panel shows some chromatin condensation whereas that in the middle, which is less typical, shows no significant chromatin condensation. Right: control embryo microinjected with injection solution at 25 min and pictured at 75 min showing condensed chromosomes on the metaphase plate. *b*, Left: two-cell embryo at 138–140 min after fertilization. The lower blastomere was microinjected at 135 min with a calcium-EGTA buffer containing 1 μM free calcium¹³ to a final concentration of 3 mM. The upper blastomere was microinjected with injection solution as a control. Chromatin condensation has occurred in the calcium-injected lower blastomere but not in the control-injected upper blastomere. Right: the same embryo 10 min later. The control (upper) blastomere has undergone normal NEB and chromosome condensation and has entered mitosis. The chromatin of the calcium-injected blastomere remains in a condensed state. This embryo subsequently underwent the second cleavage division with chromatin segregating into each of the four daughter cells. *c*, Time course of chromatin condensation after microinjection of 2.5 pl (0.5% egg volume) of 10 μM $\text{Ins}(1,4,5)\text{P}_3$ 50 min after fertilization. Left: 5 s after injection; middle: 16 s; right: 27 s.

Methods. Eggs were incubated with 20 μM Hoechst 33342 (ref. 10) in ASW (450 mM NaCl, 10 mM KCl, 11 mM CaCl_2 , 50 mM MgCl_2 , 2.5 mM NaHCO_3 , 0.5 mM EDTA, pH 8.0) and fertilized 10 min later. Fertilized eggs held on glass coverslips with polylysine were viewed with a Nikon Optiphot fluorescence microscope and microinjected using short, high pressure pulses¹³. Injection volumes ranged from 1.5–2.5 pl (0.3–0.5% egg volume) for $\text{Ins}(1,4,5)\text{P}_3$ to 20–50 pl (2–5% egg volume) for calcium-EGTA buffers. The injection solution was 0.5 M KCl, 100 μM EGTA, 20 mM PIPES, pH 6.7; this solution does not significantly alter intracellular pH¹³. Concentration of EGTA and BAPTA buffers in the pipette was 50–80 mM and of $\text{Ins}(1,4,5)\text{P}_3$ was 10 μM . The temperature was 16 °C.

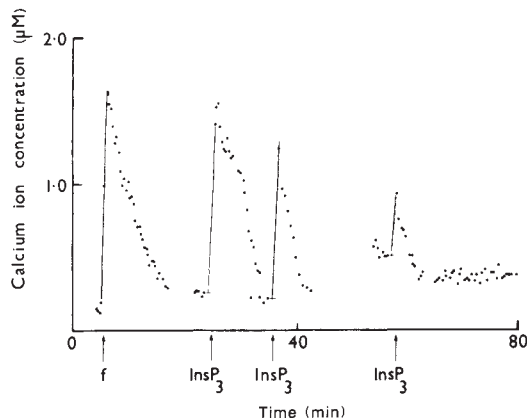


Fig. 2 Transient increases in Ca_i induced by microinjection of $\text{Ins}(1,4,5)\text{P}_3$ at various times during the first cell cycle. The change in Ca_i after fertilization in normal cells (*f*) and the results of three microinjection experiments on separate eggs are shown. The size and duration of changes in Ca_i induced by microinjecting 10 μM (5 pl, final concentration 30–50 nM) $\text{Ins}(1,4,5)\text{P}_3$ 20, 30 and 50 min after fertilization are similar. The eggs had been microinjected before fertilization with the calcium-sensitive indicator dye fura 2 (ref. 27) to a final concentration of 10 μM . This dye was used to calculate the free calcium concentration by the ratio method²⁷ using an apparent affinity constant of fura 2 for Ca^{2+} of 777 nM, and calibrating the system as described⁶.

BAPTA (1,2-bis(2-aminophenoxy)ethane- N,N,N',N' -tetraacetic acid)¹¹. NEB was prevented in all eggs microinjected 15–25 min after fertilization by a cytoplasmic BAPTA concentration of 2.3 mM (half maximal effect was observed with 1.8 mM BAPTA). In contrast TPEN (N,N,N',N' -tetrakis(2-pyridylmethyl)-ethylenediamine), a permeable chelator specific for heavy metals¹², had no effect on the eggs when added to the surrounding sea water at a concentration of 50 μM . The fact that we saw a small degree of chromatin condensation in BAPTA-injected eggs (Fig. 1*a*) indicates that the eggs are arrested at the point of normal NEB in the cell cycle. These results are consistent with the hypothesis that an increase in Ca_i leads to chromatin condensation.

In a contrasting experiment, we microinjected calcium-EGTA buffers containing 1 μM free calcium and found that this

induced premature NEB and chromatin condensation (Fig. 1*b* and Table 1). This demonstrates that elevated cytoplasmic calcium concentrations can induce the changes that normally occur immediately before entry into mitosis.

One disadvantage of calcium-EGTA buffers is that they produce a prolonged increase of Ca_i in sea urchin eggs¹³, whereas physiological changes in Ca_i are usually short lived^{14,15}. If the source of Ca_i during the cell cycle is an intracellular store, then microinjecting the calcium-mobilizing second messenger $\text{Ins}(1,4,5)\text{P}_3$ might produce a transient change in Ca_i that is closer to the changes observed in normal cells. We found that $\text{Ins}(1,4,5)\text{P}_3$ injected 50–60 min after fertilization induced a rapid condensation of the chromosomes and dissolution of the nuclear envelope (Fig. 1*c* and Table 1). It appeared to act by releasing calcium, because inclusion of the calcium chelator

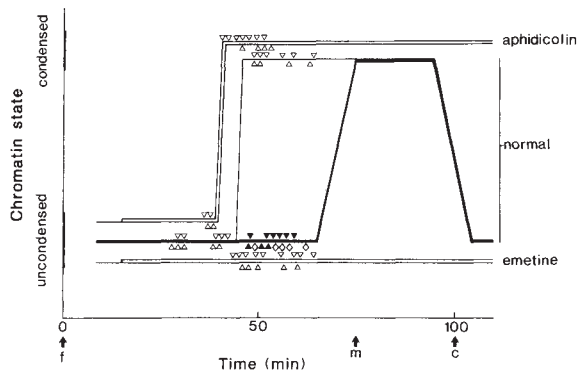


Fig. 3 Response of chromatin to $\text{Ins}(1,4,5)\text{P}_3$ during the first cell cycle and the effects of protein- and DNA-synthesis inhibitors. Each symbol corresponds to a microinjection experiment on a single egg. Variable cell cycle times were normalized to cleavage (c) at 100 min; (M) indicates the entry into mitosis. The solid line shows the normal pattern of chromatin condensation during the first cell cycle. Chromatin condensation in response to microinjection of $\text{Ins}(1,4,5)\text{P}_3$ ($10\ \mu\text{M}$; 5 pl, ∇ , Δ) first occurs 46 min after fertilization (f). Injection of solution ($\diamond$) or $\text{Ins}(1,4,5)\text{P}_3$ solutions containing 100 mM EGTA (∇ , Δ) has no effect. Treatment of eggs at 15 min after fertilization with the protein synthesis inhibitor emetine ($100\ \mu\text{M}$ in ASW; lower double line) prevents NEB and chromatin condensation even following $\text{Ins}(1,4,5)\text{P}_3$ microinjection (∇ , Δ). Eggs treated with the DNA-synthesis inhibitor aphidicolin ($100\ \mu\text{M}$ in ASW; upper double line) undergo normal NEB, but show premature chromatin condensation in response to $\text{Ins}(1,4,5)\text{P}_3$ (∇ , Δ). Uninjected aphidicolin-treated embryos undergo NEB and chromosome condensation at the same time as controls (not shown). In control experiments using the DNA stain H 33342, we found that DNA synthesis occurs in single eggs from 35 to 55 min after fertilization and that treatment with aphidicolin inhibits DNA synthesis.

EGTA in the injected solution prevented the effect (Table 1). The premature NEB and chromatin condensation induced by $\text{Ins}(1,4,5)\text{P}_3$ did not lead to a shortening of the first cell cycle: chromatin segregation and cell cleavage occurred at the same time as in control eggs (Table 1). In other words $\text{Ins}(1,4,5)\text{P}_3$ did not advance the cell cycle clock.

Microinjection of $\text{Ins}(1,4,5)\text{P}_3$ within 20 min of the time of normal NEB always resulted in rapid chromatin condensation and NEB, whereas injections before this time had no such effect (Fig. 3). One possible explanation for this discrepancy would be that $\text{Ins}(1,4,5)\text{P}_3$ fails to induce calcium release when injected into eggs very soon after fertilization. We know that this is not the case, however, as we have measured increases in Ca_i in response to $\text{Ins}(1,4,5)\text{P}_3$ injection even at these early points in the cell cycle when the nucleus is unresponsive to $\text{Ins}(1,4,5)\text{P}_3$ (Fig. 2). The non-responsiveness of the nucleus is not, in fact, surprising, as the pronucleus is clearly insensitive to the large increase in Ca_i that accompanies fertilization^{6,13}. There must therefore be a mechanism that sensitizes the nucleus to Ca_i . The observation that inhibition of protein synthesis with emetine prevents NEB and chromosome condensation¹⁶ suggested that the nucleus might also be under the control of newly synthesized protein.

Chromosome condensation in embryos of the sea urchin *Strongylocentrotus purpuratus* is sensitive to emetine until 35 min after fertilization^{16,17}. We have reproduced these results in *L. pictus*. Treatment of embryos with $100\ \mu\text{M}$ emetine up to 30 min after fertilization prevents NEB (which normally occurs 35 min later) but does not interfere with DNA synthesis (our unpublished observations and ref. 17). In contrast, treatment with emetine 40 min after fertilization does not block the first cell cycle; instead, NEB is prevented in both sister blastomeres in the second cell cycle (our observations using *L. pictus* and ref.

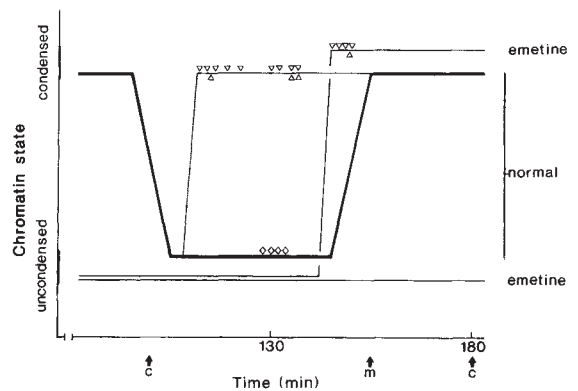


Fig. 4 Response of chromatin to $\text{Ins}(1,4,5)\text{P}_3$ during the second cell cycle. In normal eggs (solid line), chromatin condensation in response to $\text{Ins}(1,4,5)\text{P}_3$ ($10\ \mu\text{M}$; ∇ , Δ) can occur immediately after the new nuclear envelope has formed. Control injections have no effect ($\diamond$). Eggs treated 75 min after fertilization with emetine ($100\ \mu\text{M}$ in ASW) fail to undergo NEB and chromatin condensation during the second cell cycle. These emetine-treated eggs do, however, show chromatin condensation in response to $\text{Ins}(1,4,5)\text{P}_3$ (∇ , Δ). Time is minutes after fertilization; (c) cleavage; (m) mitosis.

16 for *S. purpuratus*). Figure 3 shows that the nucleus becomes sensitive to $\text{Ins}(1,4,5)\text{P}_3$ injection 45 min after fertilization, close to the time at which the critical protein synthesis blocked by emetine must have occurred. Moreover, our observations (Fig. 3) indicate that the nuclei of emetine-treated embryos are insensitive to $\text{Ins}(1,4,5)\text{P}_3$ even though emetine did not prevent $\text{Ins}(1,4,5)\text{P}_3$ -induced changes in Ca_i (data not shown). In contrast, embryos treated with the DNA polymerase inhibitor aphidicolin have the same pattern of sensitivity to $\text{Ins}(1,4,5)\text{P}_3$ as control eggs. These experiments demonstrate that protein synthesized by 40–45 min after fertilization makes the nucleus sensitive to Ca_i . It is also evident that chromatin condensation induced by an increase in Ca_i can occur well before the completion of DNA synthesis (Fig. 3).

There are two classes of protein synthesized after fertilization. One class is specific to the early, rapid divisions of the sea urchin embryo and persists until the mid-blastula stage¹⁸. The other class, the cyclins, is destroyed in each cell cycle at the time of mitosis and is thought to be involved in cell cycle control^{19,20}. Addition of protein synthesis inhibitors immediately after mitosis in the first cell cycle prevents the reappearance of cyclin¹⁹. Figure 4 illustrates that emetine added during this second sensitive period blocks NEB, presumably by preventing the resynthesis of cyclin¹⁹. In contrast to our findings in the first cell cycle, the emetine block during the second cell cycle can be relieved by increasing Ca_i using $\text{Ins}(1,4,5)\text{P}_3$. Figure 4 also shows that the nucleus lacks an $\text{Ins}(1,4,5)\text{P}_3$ -insensitive period during the normal second cell cycle. The nucleus, once sensitized to Ca_i , remains sensitive throughout subsequent cell cycles. These results demonstrate that the protein that initially confers calcium-sensitivity on the nucleus is not a cyclin and strongly suggest that the effect of preventing cyclin synthesis during the second cell cycle is to prevent the change in Ca_i responsible for NEB. Our observation (Karl Swann and M.W., unpublished data) that $100\ \mu\text{M}$ emetine abolishes the cell cycle-associated changes in calcium in early embryos supports this inference. We identify two distinct elements in the control of NEB by proteins in sea urchin embryos. One element makes the nucleus sensitive to Ca_i , and this effect is not lost in subsequent cell cycles. The other is a cyclin that is part of the mechanism that generates the change in Ca_i at NEB, probably by stimulating phosphoinositide metabolism.

Our findings show that the increase in calcium concentration associated with chromosome condensation and NEB is the

stimulus for the breakdown of the nuclear envelope. They indicate that protein synthesis at the onset of embryonic development primes the nucleus to respond to calcium, but that in subsequent cell cycles, alterations in Ca_i alone determine the timing of NEB. Thus they provide an illustration of how a common cytoplasmic signal may have different effects at different times in the same cell. We do not know the identity of the protein or proteins that prime the nucleus; they may resemble the protein kinases that mediate NEB in *Xenopus* and starfish oocytes²¹⁻²⁴, although the details of the mechanism presumably differ because calcium antagonizes chromatin condensation in *Xenopus* embryos and extracts^{25,26}. All our observations point to the phosphoinositide messenger $Ins(1,4,5)P_3$ as an important component of the mechanism of cell cycle regulation in sea urchin eggs.

This work was supported by funds from the Wellcome Trust, the Nuffield Foundation and the Dale Fund of the Physiological Society. We thank David Attwell for his advice.

Received 22 November 1987; accepted 15 February 1988

1. Gurdon, J. B. *J. Embryol. exp. Morph.* **20**, 401-414 (1968).
2. Rao, P. N. & Johnson, R. T. *Nature* **225**, 159-164 (1970).
3. Masui, H. & Markert, C. L. *J. exp. Zool.* **117**, 129-146 (1971).
4. Baker, P. F. & Metcalfe, J. C. *Br. med. Bull.* **42**, 341 (1986).
5. Poenie, M. & Steinhardt, R. A. in *Calcium and Cellular Function* Vol. 7 (ed. Cheung, M.) 133-157 (Academic, New York, 1987).
6. Poenie, M., Alderton, J., Tsien, R. Y. & Steinhardt, R. A. *Nature* **315**, 147-149 (1985).
7. Steinhardt, R. A. & Alderton, J. *Nature* **332**, 364-366 (1988).
8. Yoshimoto, Y., Iwamatsu, T. & Hiramoto, Y. *Biomed. Res.* **6**, 387-394 (1985).
9. Hafner, M. & Petzelt, C. *Nature* **330**, 264-266 (1987).
10. Arndt-Jovin, D. J. & Jovin, T. M. *Histochem. Cytochem.* **25**, 585-589 (1977).
11. Tsien, R. Y. *Biochemistry* **19**, 2,396-2,404 (1980).
12. Arslan, P., Di Virgilio, F., Beltrame, M., Tsien, R. Y. & Pozzan, T. *J. biol. Chem.* **260**, 2,719-2,727 (1985).
13. Swann, K. & Whitaker, M. J. *J. Cell Biol.* **103**, 2,333-2,342 (1986).
14. Irvine, R. F. *Br. med. Bull.* **42**, 369-374 (1986).
15. Woods, N. M., Cuthbertson, K. S. R. & Cobbold, P. H. *Nature* **319**, 600-602 (1986).
16. Wagenaar, E. B. & Mazia, D. in *Cell Reproduction* (eds Dirksen, E., Prescott, D. & Fox, C. F.) 539-545 (Academic, New York, 1978).
17. Wagenaar, E. B. *Expl. Cell Res.* **144**, 393-403 (1983).
18. Grainger, J. L., von Brunn, A. & Winkler, M. M. *Dev. Biol.* **114**, 403-415 (1986).
19. Evans, T., Rosenthal, E. T., Youngblom, J., Distel, D. & Hunt, T. *Cell* **33**, 389-396 (1983).
20. Swenson, K. I., Farrell, K. M. & Ruderman, J. V. *Cell* **47**, 861-870 (1986).
21. Miake-Lye, R., Newport, J. & Kirschner, M. W. *J. Cell Biol.* **97**, 81-91 (1983).
22. Miake-Lye, R. & Kirschner, M. W. *Cell* **41**, 165-175 (1985).
23. Suprynowicz, F. A. & Gerace, L. J. *Cell Biol.* **103**, 2,073-2,081 (1986).
24. Picard, A. *et al.* *Dev. Growth Different.* **29**, 93-103 (1987).
25. Huthison, C. J., Cox, R., Drepaal, R. S., Gomperts, B. D. & Ford, C. C. *EMBO J.* **6**, 2,003-2,010 (1987).
26. Newport, J. W. & Kirschner, M. W. *Cell* **37**, 731-742 (1984).
27. Gryniewicz, G., Poenie, M. & Tsien, R. Y. *J. biol. Chem.* **260**, 3,440-3,450 (1985).

Allele-specific and *Mutator*-associated instability at the *Rp1* disease-resistance locus of maize

Jeffrey L. Bennetzen*, Min-Min Qin†, Stephen Ingels† & Albert H. Ellingboe†

* Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA

† Departments of Plant Pathology and Genetics, University of Wisconsin, Madison, Wisconsin 53706, USA

The *Rp1* locus of maize determines resistance to the fungal rust pathogen *Puccinia sorghi*. In an attempt to isolate insertion mutations at *Rp1* with the *Mutator* transposable element system, we found that *Rp1* inactivation was common both in lines with *Mutator* activity and in controls lacking any known transposable element activity. Some alleles of *Rp1* are endogenously inactivated as frequently as one in five-hundred, whereas other alleles are over twenty times more stable. In a standard background, no mutations of the *Rp1*^F allele were detected in 7,339 progeny screened. In a

Table 1 Allele- and background-specific instability at *Rp1*

<i>Rp1</i> allele	Background	Number of seedlings screened	Number of susceptible seedlings	Number of susceptible seedlings with progeny tested*	Number of susceptibles based on progeny test
<i>Rp1</i> ^D	Standard (R168)	11,304	0	0	0
<i>Rp1</i> ^E	Standard (R168)	20,729	2	2	1
<i>Rp1</i> ^F	Standard (R168)	7,339	1	1	0
<i>Rp1</i> ^G	Standard (R168)	3,917	7	5	5
<i>Rp1</i> ^K	Standard (R168)	18,152	4	0	0
<i>Rp1</i> ^L	Standard (R168)	1,069	2	2	2
<i>Rp1</i> ^F	<i>Mutator</i>	35,356	38	27	25†
<i>Rp1</i> ^K	<i>Mutator</i>	11,752	9	6	0

* Some susceptible seedlings did not yield progeny.

† The progeny of two of the original seedlings scored as susceptibles (A-8, A-43) segregated intermediate *Rp1*^F resistance.

Mutator background, 27 *Rp1*^F inactivations were isolated from 35,356 tested seedlings. Although most *Rp1*^F mutant seedlings lost resistance to all of the *P. sorghi* isolates tested, several retained resistance to a subset of rust biotypes. Some of the resistance specificity profiles obtained in these mutants are unlike any previously identified *Rp1* alleles. These data indicate that the *Rp1* locus is composed of multiple resistance factors with variable intrinsic stabilities.

Scores of genetic loci that determine complete or partial resistance to specific microbial pathogens have been identified in a variety of plant species^{1,2}. Many of these disease-resistance genes are dominant and specify resistance to particular races of a given pathogen. The genetics of this race-specific resistance indicate that the disease-resistance gene in the plant recognizes a single dominant gene product or process in the pathogen, leading to an incompatible host-pathogen interaction, namely, resistance^{3,4}. The molecular mechanism of this gene-for-gene resistance interaction has not been elucidated.

One approach to understanding the race-specific resistance phenomenon would be to molecularly clone and determine the nature of the interacting plant and pathogen genes. Initially we attempted to clone a plant disease-resistance gene by the technique of transposon tagging. We chose the *Mutator* system⁵ of maize as the tagging module because of its uniquely high mutagenic activity⁶ and the availability of the cloned transposable element⁷. We set out to clone the best studied gene-for-gene resistance locus in maize, *Rp1*, a gene specifying resistance to the leaf rust fungus *Puccinia sorghi*⁸⁻¹⁰.

Several alleles have been defined at the *Rp1* locus by the unique spectrum of *P. sorghi* race-specific resistance they provided^{9,10}. Alleles *Rp1*^A through *Rp1*^N were backcrossed for ten generations into the rust-susceptible background R168. We have sibling-crossed these near-isogenic maize lines and have identified and amplified *Rp1* homozygous stocks for each allele. Although most of the *P. sorghi* races identified by Hooker *et al.*⁹ have since been lost, we have isolated several different *P. sorghi* biotypes that differentiate most *Rp1* alleles.

Maize lines carrying alleles *Rp1*^F or *Rp1*^K were separately crossed to *Mutator* lines and then sibling crossed. *Rp1*^F or *Rp1*^K homozygous lines were identified and tested for *Mutator* mutagenic activity⁵. *Rp1* homozygous lines with *Mutator* activity were then crossed as females to a male *rp1/rp1* line (Fig. 1). The kernels from each ear derived from these crosses were planted separately in the greenhouse. Resultant seedlings were inoculated with *P. sorghi* spores from a field culture consisting of eight distinguishable rust biotypes. Out of 35,356 *Rp1*^F/*rp1* progeny screened, 38 susceptible seedlings were identified (Table 1). Twenty-seven of these putative *Rp1*^F mutations survived and were either self-crossed or backcrossed to the *rp1* tester. On rescreening with *P. sorghi*, twenty-five of the twenty-seven putative backcrossed or selfed F₁ plants yielded only

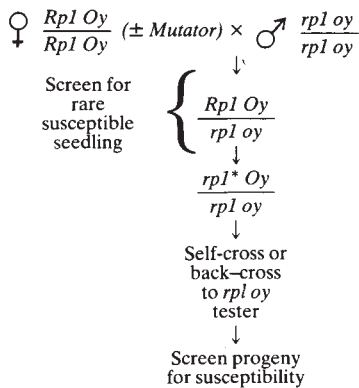


Fig. 1 Crossing programme to isolate mutations at the *Rpl* locus. Homozygous female *Rpl* plants, with or without *Mutator* activity, were crossed to homozygous male *rpl* (rust-susceptible) plants that also carried the recessive, flanking *oy* marker. The F_1 progeny of this cross were tested for rust-resistance by spraying 10-day-old seedlings with a suspension of *P. sorghi* spores. The *P. sorghi* inoculum used was a field isolate consisting of at least eight differentiable rust races. Susceptible seedlings were self-crossed or backcrossed to the *rpl/rpl* tester. Progeny of this cross were then screened, as above, for resistance to the *P. sorghi* field isolate. *Oy* and *oy* are alleles of the *oil yellow* locus which condition dominant green and recessive oil yellow plant colour, respectively. *rpl** is a mutant allele of *Rpl*.

susceptible progeny. Offspring of the other two plants segregated for intermediate resistance to our standard field culture.

Screening 11,752 progeny of Rpl^K/Rpl^K *Mutator* stocks crossed to the *rpl/rpl* tester yielded nine susceptible seedlings. Six of these seedlings survived and were crossed and rescreened as described for Rpl^F . The progeny of all six putative Rpl^K mutants tested for susceptibility segregated complete Rpl^K resistance. We believe these escapes are due to some physiological deficiency (such as root rot) in the original seedling tests.

At the same time that the Rpl^F and Rpl^K *Mutator* lines were test crossed and scored for *Rpl* mutations, we followed the identical crossing and screening programme for several *Rpl* alleles in the rust-susceptible R168 background. Surprisingly, three alleles of *Rpl* (Rpl^E , Rpl^G , Rpl^L) proved to be intrinsically unstable (Table 1). As our sample populations were small in these control screenings, we could not guarantee that any of the other alleles screened (Rpl^D , Rpl^F , Rpl^K) were more stable than 10^{-4} . The differences in *Rpl* inactivation rates between the highly unstable alleles Rpl^G and Rpl^L (5–7/3,917 and 2/1,069, respectively) and the more stable alleles Rpl^D , Rpl^E and Rpl^F , (0/11,304, 1/20,729, and 0/7339, respectively) is statistically significant ($P < 0.01$). Recently, similar results have been reported by Pryor indicating endogenous, allele-specific instability at *Rpl* (ref. 11).

In further crosses to test the subsequent stability of the *Mutator*-associated Rpl^F mutations, we have not found any reactivation of Rpl^F in over 13,000 seedlings screened. This indicates that these endogenous mutations revert relatively rarely in germinal tissues, as is also true of many *Mutator*-induced mutations¹².

Several of the confirmed Rpl^F mutations isolated from a *Mutator* line were separately tested for susceptibility to isolated rust biotypes (Table 2). Twenty-one Rpl^F mutations were equally susceptible to all six biotypes inoculated, whereas six mutants retained resistance to a subset of rust biotypes. The patterns observed were consistent with that expected in a gene-for-gene system^{3,4}. For example, mutant B-18 was susceptible to rust isolate 1-2 but resistant to isolate 1-3, and mutant D-3 was resistant to isolate 1-2 but susceptible to isolate 1-3 (Table 2). All but one (Rpl^B) of the unique, previously identified *Rpl*

Table 2 Reaction of Rpl^F -derived susceptibles from a *Mutator* background to six rust isolates

Designation of Rpl^F mutant	<i>P. sorghi</i> biotype/reaction*					
	Biotype 1-2	Biotype 1-3	Biotype 1-4	Biotype 1-6	Biotype 1-7	Biotype 1-8
A-2	S	S	S	S	S	S
A-8†	S	R	R	S	R	R
A-13	S	S	S	S	S	S
A-65	S	S	S	S	S	S
B-3	R	R	R	S	R	R
B-18	S	R	S	S	S	S
D-3†	R	S	S	S	S	S
D-22 ^a	S	S	S	S	S	S
D-25†	S	S	R	S	R	R
Rpl^F	R	R	R	R	R	R

* Resistant, R; susceptible, S.

† Progeny of these mutants were from backcrosses to the *rpl/rpl* tester; all others were from self-crosses. Eight or more seedlings of the selfed Rpl^F mutants, and twelve or more seedlings of the backcrossed Rpl^F mutant seedlings, were inoculated with each isolate. If all progeny were susceptible, they were classified S. If one or more of the seedling progeny were resistant to that rust isolate, the classification was R.

alleles⁹ were also screened against these six rust biotypes. None of these *Rpl* alleles had a biotype specificity profile identical to any of the Rpl^F partial-spectrum mutants (A-8, B-3, B-18, D-3, or D-25) (data not shown).

Several other rust-resistance genes (*Rp5*, *Rp6*, and *Rpp9* map near *Rpl* (ref. 10). If the distal region of the short arm of chromosome 10 in maize is composed of multiple *Rp* genes of similar sequence, the instability of the locus could be due to unequal crossovers¹³ that generate deletions on one homologue and duplications on the other. The relative instability of various alleles could then be due to their position in the *Rp* gene array. Central *Rp* genes would be lost by unequal crossing-over more commonly than would genes at the ends of the cluster.

The retention of resistance to a subset of rust races by some Rpl^F mutants suggests that more than one resistance component can be active in a single *Rpl*-containing cell. At Rpl^F , six classes of mutants have been detected, indicating at least four separable genetic determinants of resistance at Rpl^F . As the five classes of Rpl^F mutants which retained resistance to some rust biotypes had resistance ranges unlike Rpl^A or Rpl^C – Rpl^N , it is possible that these *Mutator*-associated Rpl^F derivatives were generated by a process different from the phenomenon which derived Rpl^A – Rpl^N from a primordial *Rpl* locus.

We thank A. Hooker and R. Wong for discussion, W. Pederson for *Rpl* seed, A. Pryor for discussion and for permission to cite his recent findings, and A. Maki, C. Carter and P. Herzmark for technical assistance. This research was supported by grants from Eli Lilly and an NSF predoctoral fellowship (S.I.).

Received 10 November 1987; accepted 4 February 1988.

- Day, P. R. *Genetics of Host-Parasite Interaction* (Freeman, San Francisco, 1974).
- Sidhu, G. S. *Sci. Prog.* **62**, 467–485 (1975).
- Flor, H. H. *Adv. Genet.* **8**, 29–54 (1956).
- Ellingboe, A. H. in *Physiological Plant Pathology, Encyclopedia of Plant Physiology*, Vol. 4 (eds Heitefuss, R. & Williams, P. H.) 761–778 (Springer, New York and Heidelberg, 1972).
- Robertson, D. S. *Mutat. Res.* **51**, 21–28 (1978).
- Robertson, D. S. & Mascia, P. N. *Mutat. Res.* **84**, 283–289 (1981).
- Bennetzen, J. L., Swanson, J., Taylor, W. C. & Freeling, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4125–4128 (1984).
- Rhoades, V. H. *Proc. natn. Acad. Sci. U.S.A.* **21**, 243–246 (1935).
- Wilkinson, D. R. & Hooker, A. L. *Phytopathology* **58**, 605–608 (1968).
- Saxena, K. M. S. & Hooker, A. L. *Proc. natn. Acad. Sci. U.S.A.* **61**, 1300–1305 (1968).
- Pryor, A. J. *Maize Genet. Coop. Newsl.* **61**, 37–38 (1987).
- Bennetzen, J. L. *Molec. cell. Biol.* **35**, 343–353 (1985).
- Sturtevant, A. H. *Genetics* **10**, 117–147 (1925).

The *GLI* gene is a member of the *Kruppel* family of zinc finger proteins

Kenneth W. Kinzler, John M. Ruppert,
Sandra H. Bigner* & Bert Vogelstein†‡

Departments of Oncology and Pharmacology, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

* Department of Pathology, Duke University Medical Center, Durham NC 27710, USA

† Oncology Research Laboratories, The Johns Hopkins Oncology Center, 424 N. Bond Street, Baltimore, MD 21231, USA.

Many studies have established that a select subset of normal cellular genes are altered in cancer by point mutations, translocations or gene amplification¹. However, the vast majority of genetic changes that occur in neoplastic cells have not yet been identified. In an attempt to identify some of these other genetic changes, we have recently isolated a gene, *GLI*, by virtue of its amplification in a human glioblastoma². Subsequently, *GLI* was found to be amplified in other human glioblastomas (ref. 3 and unpublished data). To understand better the role of *GLI* in human neoplasia, we have now cloned the *GLI* complementary DNA (cDNA) and determined its nucleotide sequence. Analysis of the predicted translation product reveals that it contains five repeats of a DNA binding consensus sequence (zinc finger) originally described in *Xenopus* Transcription Factor III A (TFIIIA)⁴. Furthermore, these zinc fingers contain sequence elements that suggest the *GLI* gene product is a member of the recently described *Kruppel* family of zinc finger proteins^{5,6}. Additional experiments demonstrate that *GLI* is an evolutionarily conserved gene that is expressed in embryonal carcinoma cells but not in most adult tissues. The link between the developmentally important *Kruppel* family of genes and *GLI* is interesting considering the similarities between developing embryonic and neoplastic tissue.

Our source of RNA for the cloning of the *GLI* cDNA was the D-259 MG cell line. This cell line was derived from a glioblastoma⁷ that had a 75-fold amplification of the *GLI* gene². A plasmid containing a 1.55 kilobase (kb) genomic *Pst*I fragment of the *GLI* gene (pKK36P1, ref. 2) detected a 4.0 kb mRNA transcript in D-259 MG cells in hybridization experiments (data not shown). The insert from pKK36P1 was used to screen an oligo(dT) primed cDNA library prepared from mRNA isolated from D-259 MG cells. The longest cDNA clone identified was λ M5E which extended 1,400 base pairs (bp) from the poly(A) tail (Figure 1a). To isolate clones representing sequences farther upstream of λ M5E, we constructed a random hexamer-primed cDNA library from D-259 MG mRNA. This library was first screened with the 5' end of λ M5E (probe A, Fig. 1a) and clone λ H31 was found to extend the farthest upstream of any of the clones isolated. In a similar manner, the 5' ends of λ H31 (probe B) and λ I2G (probe C) were used to isolate clones λ I2G and λ J36, respectively (Fig. 1a). All four of these cDNA clones were representative of the major *GLI* transcript as determined by S₁ nuclease or RNase A protection analysis. Taken together these four cDNA clones encompass 3.6 kb (not including the Poly(A) tail) of the 4.0 kb *GLI* transcript. Isolating sequences that extended farther upstream than λ I2G proved to be difficult because nine out of the sixteen clones analysed had their 5' terminus within 120 bp of the 5' end of λ I2G and the other seven transcripts had varied 5' ends and did not represent major *GLI* transcripts. This heterogeneity at the 5' end may be associated with a genomic rearrangement that occurs upstream of λ J36 in D-259 MG cells (J. Nigro *et al.* unpublished observations).

Sequence analysis of these cDNA clones revealed a 3,318 bp long open reading frame preceded by 78 bp of non-translated sequence (Fig. 1b). We chose the methionine codon at nucleotide 1 as the predicted start site of translation for two reasons. First, the ATG is preceded by an in-frame stop codon at nucleotide -9. Second, the sequence surrounding this methionine codon (ACGCCATGT) fits the consensus sequence for a eukaryotic translation start site (CC^A/GCCATGG) at four out of six positions⁸. A more favourable site for initiation does not occur until the seventeenth methionine codon at nucleotide 2,092. The open reading frame is terminated by a stop codon at nucleotide 3,319 and followed by a cluster of four in-frame stop codons at nucleotides 3,469, 3,481, 3,484 and 3,487. This cluster of stop codons is followed by a poly(A) addition signal (AATAAA) at nucleotide 3,504 and a poly(A) tail at nucleotide 3523.

The open reading frame could encode for an 1,106 residue protein with a calculated molecular mass of 118 Kdalton. The predicted *GLI* protein sequence was compared with protein sequences contained in the Protein Identification Resource data bank (R12.0, May 1987). The only protein found to have a significant degree of identity with *GLI* was TFIIIA, which displayed 34% identity over 118 residues. The region of homology between TFIIIA and *GLI* corresponds to the zinc finger region of TFIIIA⁴. Examination of the *GLI* sequence reveals that it has five repeats of the zinc finger motif, (F/YX₂CX₂4CX₃FX₂LX₂HX₃4HX₅) between residues 235 and 393 (Fig. 1c, ref. 9). The essential cysteine and histidine residues are invariant and the finger sizes vary from 30–33 residues. Comparison of the *GLI* finger sequence with other zinc finger containing proteins indicates that it has consensus features representative of the *Kruppel* family of zinc finger proteins^{5,6}. The *Kruppel* gene is a *Drosophila* segmentation gene of the gap class¹⁰. It has five zinc fingers which are noteworthy for their conservation of the sequence HTGEKP between the last histidine of one finger and the first cysteine of the next finger (H-C link)^{5,6,11}. Other members of the family have been identified by their hybridization to DNA probes encoding zinc fingers with the H-C link^{5,6}. The H-C link is conserved in the second, third and fourth *GLI* finger and to a lesser extent in the first *GLI* finger (Fig. 1c). Furthermore, when the *GLI* zinc finger sequences are used as a hybridization probe, they detect several related human zinc finger sequences with highly conserved H-C links (Ruppert *et al.*, manuscript submitted). Hence *GLI* is the first human member of the *Kruppel* family to be identified and is the prototype of a family of human genes.

On further examination of the sequence we found seventeen serine residues and three threonine residues in a context which makes them suitable substrates for various protein kinases¹². Analysis of the *GLI* protein also revealed that the zinc fingers are bounded by two hydrophilic regions (residues 216–230 and residues 412–428). There is no obvious nuclear localization sequence although there is an arginine/lysine rich region (five out of ten residues) at residues 72–81.

Although the *GLI* gene is a member of the *Kruppel* family, it is not the homologue of any previously described *Kruppel* family gene because the similarity does not extend beyond the zinc finger domains. To determine whether a *GLI* homologue exists in other species, we used various regions of the *GLI* gene as probes for Southern blot analysis of 'zoo' blots under reduced stringency. Our zoo contained DNA from humans, rats, mice, chickens and flies. A probe for the region preceding and including the first zinc finger (nucleotides 40–795) detected one or two restriction fragments in each species tested (Fig. 2a). As expected, a cDNA for the zinc finger domain (nucleotides 796–1,189) detected numerous related sequences in all species tested when hybridized at low stringency (data not shown). When more stringent conditions were used, a smaller number of strongly hybridizing restriction fragments were detected in all five species (Fig. 2b). Finally, when a cDNA probe for the region downstream of the zinc finger domain (nucleotides 1,177–2,348) was

‡ To whom correspondence should be addressed.

b

ATG ACC CCA CCA CCA ATC AGT AGC TAT GGC GAG CCC TGC TGT CTC CGG CCC CTC CCC AGT CAG GGG GCC CCC AGT GTG GGG ACA GAA GGA	bp12
Met Thr Pro Pro Pro Ile Ser Ser Tyr Gly Glu Pro Cys Cys Leu Arg Pro Leu Pro Ser Glu Gly Ala Pro Ser Val Gly Thr Glu Gly	AA4
CTG TCT GGC CGC CCC TTC TGC CAC GAT AAC CTC ATG TGC GGC CCC CAC AGT TAT GGG CCA GCC AGA GAG ACC AAC AGC TGC ACC GAG	bp102
Leu Ser Gly Pro Pro Phe Cys His Glu Ala Asn Leu Met Ser Gly Pro His Ser Tyr Gly Pro Ala Arg Glu Thr Asn Ser Cys Thr Leu Asp	AA34
GGC CCA CTC TTT TCT TCT CCC CGG AGT GCA GTC AAG TTG ACC AAG AAG CGG GCA CTG TCC ATC TCA CTT CTG TCG GAT GCC AGC CTG GAC	bp192
Pro Leu Phe Ser Ser Pro Arg Ser Ala Val Lys Leu Thr Lys Lys Arg Ala Leu Ser Ile Ser Pro Leu Ser Asp Ala Ser Leu Asp	AA64
CTG CAG ACG GTT ATC CGC ACC TCA CCC AGC TCC CTC GTA GCT TTC ATC AAC TCG CGA TGC ACA TCT CCA GGA GGC TCC TAC GGT CAT CTC	bp282
Leu Glu Thr Val Ile Arg Thr Ser Pro Ser Ser Leu Val Ala Phe Ile Asn Ser Arg Cys Thr Ser Pro Gly Gly Ser Tyr Gly His Leu	AA94
TCC ATT GGC ACC ATG AGC CCA TCT CTG GGA TTC CCA GCC CAG ATG AAT CAC CAA AAA GGG CCC TCG CTT TCC TTT GGG GTC CAG CTT TGT	bp372
Ser Ile Gly Thr Met Ser Pro Ser Leu Gly Phe Pro Ala Glu Met Asn His Glu Lys Gly Pro Ser Pro Ser Phe Gly Val Glu Pro Cys	AA124
GGT CCC CAT GAC TCT GGC GGT GGG ATG ATC CCA CAT CCT CAG TCC CGG GGA CCC TTC CCA ACT TGC CAG CTG AAG TCT GAG TGC GAC	bp462
Gly Pro His Asp Ser Ala Arg Gly Gly Met Ile Pro His Pro Glu Ser Arg Gly Pro Phe Pro Thr Cys Glu Leu Lys Ser Glu Leu Asp	AA154
ATG CTG GTT GGC AAG TGC CGG GAG GAA CCC TTG GAA GGT GAT ATG TCC AGC CCC AAC TCC ACA GGC ATA CAG GAT CCC CTG TTG GGG ATG	bp552
Met Leu Val Gly Lys Cys Arg Glu Glu Pro Leu Glu Gly Asp Met Ser Ser Pro Asn Ser Thr Gly Ile Glu Asp Pro Leu Leu Gly Met	AA184
CTG GAT GGG CGG GAG CAG CTC GAG AAG GAG GAG AAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG	bp642
Leu Asp Gly Arg Glu Asp Leu Glu Arg Glu Lys Arg Glu Pro Glu Ser Val Tyr Glu Thr Asp Cys Arg Trp Asp Gly Cys Ser Glu	AA214
GAA TTT GAC TCC CAA GAG CAG CTG GTG CAC CAC ATC AAC AGC GAG CAC ATC CAC GGG GAG CGG AAG GAG TTC GTG TGC CAC TGG GGG GGC	bp732
Glu Phe Asp Ser Glu Glu Thr Val His His Ile Asn Ser Glu His Ile His Gly Glu Arg Lys Glu Phe Val Cys His Thr Ile Arg	AA244
TGC TCC AGG GAG CTG AAG CCC TTC AAA GCC CAG TAC ATG CTG GTT CAC ATG CGC AGA CAC ACT GGC GAG AAG CCA CAC AAG TGC ACG	bp822
Cys Ser Glu Arg Lys Leu Arg Pro Phe Lys Ala Glu Tyr Met Leu Val Val His Met Arg Arg His Thr Gly Glu Lys Pro Lys Cys Thr	AA274
TTT GAA GGG TGC CGG AAG TCA TAC TCA CGC CTC GAA AAC CTG AAG ACG CAC CTG CGG TCA CAC ACG GGT GAG AAG CCA TAC ATG TGT GAG	bp912
Phe Glu Phe Gly Cys Arg Lys Ser Tyr Ser Arg Leu Glu Asn Leu Lys Thr His Leu Arg Ser His Thr Gly Val Lys Pro Tyr Met Cys	AA304
CAC GAG GGC TGC AGT AAA GCC TTC AGC AAT GCC AGT CAG GGC AAG CAC CAG AAT CGG ACC CAT TCC AAT GAG AAG CCG TAT GTA TGT	bp1002
His Glu Gly Cys Ser Lys Ala Phe Ser Asn Ala Ser Asp Arg Ala Lys His Glu Asn Arg Thr His Ser Asn Glu Lys Pro Tyr Val Cys	AA334
AAG CTC CCT GGC TGC ACC AAA CGC TAT ACA GAT CCT AGC TCG CTG CAG AAA CAT GTC AAG ACA GTG CAT GGT CCT GAC GCC CAT GTG ACC	bp1092
Leu Cys Thr Thr Lys Arg Thr Thr Asn Pro Ser Ser Leu Arg Lys His Val Lys Thr Val His Gly Pro Asp Ala His Val Thr	AA364
AAA CGG CAG CBT GGG GAT GGC CCC CTG CCT CGG GCA CCA TCC ATT TCT ACA GTG GAG CCC AAG AGG GAG CGG GAA GGA GGT CCC ATC AGG	bp1182
Lys Arg His Arg Lys Asp Gly Pro Leu Pro Arg Ala Pro Ser Ile Ser Thr Val Glu Pro Lys Arg Glu Arg Glu Gly Pro Ile Arg	AA394
GAG GAA AGC AGA CTG ACT GTG CCA GAG GGT GCC ATG AAG CCA CAG CCA AGC CCT GGG GCC CAG TCC TGC AGC AGT GAC CAC TCC CGG	bp1272
Glu Glu Phe Ser Arg Leu Thr Val Pro Glu Gly Ala Met Lys Pro Glu Pro Ser Pro Gly Ala Glu Ser Ser Cys Ser Ser Asp His Ser Pro	AA424
GCA GGG AGT GCA GCC AAT ACA GAG AGT GGT GTG GAA ATG ACT GGC AAT GCA GGG GGC AGC ACT GAA GAC CTC TCC AGC TTG GAG GGA	bp1362
Ala Gly Ser Ala Ala Asn Thr Asp Ser Gly Val Glu Met Thr Gly Asn Ala Gly Gly Ser Thr Glu Asp Leu Ser Ser Leu Asp Glu Gly	AA454
CCT TGC ATT GCT GGC ACT GGT CTG TCC ACT CTT CGC CGC CTT GAG AAC CTC AGG CTG GAC CAG TCA CAT CAA CTC CGG CCA ATA GGG ACC	bp1452
Val Thr Ile Ala Gly Thr Gly Leu Ser Thr Leu Arg Arg Leu Glu Asn Leu Arg Leu Ser Glu His Glu Leu Arg Pro Ile Gly Thr	AA484
CGG GGT CTC AAA CTC CCC AGC TTG TCC CAC ACC GGT ACC ACT GTG TCC CGC CGC GTG GGC CCC CCA GTC TCT CTT GAA CGC CGC AGC AGC	bp1542
Arg Gly Leu Lys Leu Pro Ser Leu Ser His Thr Gly Thr Thr Val Ser Arg Arg Val Gly Pro Pro Val Ser Leu Glu Arg Ser Ser	AA514
AGC TCC AGC AGC ATC AGC TCT GCT TAT ACT GTC AGC CGC CGC TCC CTC GGC TCT CTT TCC CCT GGC TCC CCA CAG AAT GGA	bp1632
Ser Ser Ser Ile Ser Ser Ala Tyr Thr Val Ser Arg Ser Ser Ser Leu Ala Ser Pro Phe Pro Pro Gly Ser Pro Glu Asn Gly	AA544
GCA TCC TCC CTG CCT GGC CTT ATG CCT GGC CAG CAC TAC CTG CTT CGG GCA AGA TAT GCT TCA GCC AGA GGG GGT GGT ACT TGC CCC ACT	bp1722
Ala Ser Ser Leu Pro Gly Leu Met Pro Ala Glu His Tyr Leu Leu Arg Ala Arg Tyr Ala Ser Ala Arg Gly Gly Thr Ser Pro Thr	AA574
GCA GCA TCC AGC CTG GAT CGG ATA GGT GGT CTT CCC ATG CCT CTT TGG AGA AGC CAG GCC GAG TAT CCA GGA TAC AAC CCC AAT GCA GGG	bp1812
Ala Ala Ser Ser Leu Asp Arg Ile Gly Gly Leu Pro Met Pro Pro Trp Arg Ser Arg Ala Glu Tyr Pro Gly Tyr Asn Pro Asn Ala Gly	AA604
GTG ACC CGG AGG GCC AGT GAC CCA GCC CAG GCT GCT GAG CAG GCT GCT CCA GCT AGA GTC CAG AGG TTC AAG AGC CTG GGC TGT GTC CAT	bp1902
Val Thr Arg Pro Arg Pro Asp Pro Ala Glu Ala Asp Arg Pro Ala Pro Ala Arg Val Glu Arg Phe Lys Ser Leu Cys Val His	AA634
ACC CCA CCC ACT GTG GCA GGG GGA GGA CAG AAC TTT GAT CCT TAC CTC CCA ACC TCT GTC TAC TCA CCA CAG CCC CCC AGC ATC ACT GAG	bp1992
Thr Pro Thr Thr Val Ala Gly Gly Glu Glu Asn Phe Asp Pro Tyr Leu Pro Thr Ser Val Tyr Ser Pro Glu Pro Glu Pro Ile Thr Glu	AA664
AAT GCT GGC ATG GAT GCT GAG GGG CTA CAG GAA GAG CCA GAA GGT GGG ACC TCC ATG GTG GGC AGT GGT CTG AAC CCC TAT ATG GAC TTC	bp2082
Asn Ala Ala Met Asp Ala Arg Gly Leu Glu Glu Glu Pro Glu Val Gly Thr Ser Met Val Gly Ser Gly Leu Asn Pro Tyr Met Asp Phe	AA694
CCA CCT ACT GAT ACT CTG GGA TAT GGG GGA CCT GAA GGG GCA GCA GCT GAC CCT TAT GGA GCG AGG GGT CCA GGC TCT CTG CTT GGT GGG	bp2172
Pro Thr	AA724
CCT GGT CCA CCC ACC AAT TAT GGC CCC AAC CCC TGT CCC CAG GGC TCA TAT CCT GAC CCC ACC CAA GAA ACA TGG GGT GAG TTC CCT	bp2262
Pro Gly Pro Thr Cys Tyr Gly Pro Lys Asn Pro Cys Pro Glu Glu Ala Ser Tyr Pro Asp Pro Thr Glu Thr Thr Thr Thr Thr Thr Thr	AA754
TCC CAC TCT GGG CTG TAC CCA GGC CCC AAG GCT CTA GGT GGA ACC TAC AGC CAG TGT CCT CCA CTT GAA CAT TAT GGA CAA GTG GAA GTC	bp2352
Ser His Ser Gly Leu Tyr Pro Gly Pro Lys Ala Leu Gly Gly Thr Tyr Ser Glu Cys Pro Arg Leu Glu His Tyr Glu Glu Val Glu Val	AA784
AAG CCA GAA CAG GGG TGC CTA GTG GGG TCT GAC TCC ACA GGC CTG GCA CCC TGC CTC AAT GCC CAC CCC AGT GAG GGG CCC CCA CAT CCA	bp2442
Lys Pro Glu Glu Gly Cys Pro Val Gly Ser Asp Ser Thr Gly Leu Ala Pro Cys Leu Asn Ala His Pro Ser Glu Gly Pro Pro His Pro	AA814
CAG CCT CTC TTT TCC CAT TAC CCC CAG CCC TCT CCT CCC CAA TAT CTC CAG TCA GGC CCC TAT ACC CAG CCA CCC CTT GAT TAT CTT CCT	bp2532
Glu Pro Leu Phe Ser His Tyr Pro Glu Pro Ser Pro Glu Thr Leu Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr	AA844
TCA GAA CCC AGG CCT GTC CAG TTT GAT TCC CCC ACC CAT TCC ACA GGG CAG CTC AAG GCT CAG CTT GTG TGT AAT TAT GTT CAA TCT	bp2622
Ser Glu Pro Arg Pro Pro Cys Leu Asp Phe Ser Pro Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr	AA874
CAA CAG GAG CTA CTG TGG GAG GGT GGG GGC AGG GAA GAT GCC CCC GGC CAG GAA CCT TCC TAC CAG AGT CCG AAG TTT CTG GGG GGT TCC	bp2712
Glu Glu Thr Leu Leu Thr Trp Glu Gly Gly Arg Glu Asp Ala Pro Ala Glu Glu Pro Ser Tyr Thr Glu Ser Pro Lys Phe Leu Gly Ser	AA904
CAG GTT AGC CCA AGC GCT GCT AAA GGT CCA GTG AAC ACA TAT GGA CCT GGC TTT GGA CCC AAC TTG CCC AAT CAC AAG TCA GGT TCC TAT	bp2802
Glu Val Ser Pro Ser Arg Ala Lys Ala Pro Val Asn Thr Tyr Gly Pro Gly Phe Gly Pro Asn Leu Pro Asn His Lys Ser Gly Ser Tyr	AA934
CCC ACC CCT TCA CCA TGC CAT GAA AAT TTT GTA GTG GGG GCA AAT AGG GCT TCA ACT AGG GCA GCA CCA CCT GSA CTT CTG CCC CCA	bp2892
Pro Thr Pro Ser Pro Cys His Glu Asn Phe Val Gly Ala Asn Arg Ala Ser His Arg Ala Ala Ala Pro Pro Arg Leu Leu Pro Pro	AA964
TTG CCC ACT TGC TAT GGG CTT CTC AAA GTG GGA GGC ACA AAC CCC AGC TGT GGT CAT CCT GAG GTG GGC AGG CTA GGA GGG GGT CCT GCC	bp2982
Leu Pro Thr Cys Tyr Gly Pro Lys Lys Val Gly Gly Thr Asn Pro Ser Cys Gly His Pro Glu Val Gly Arg Leu Gly Thr Gly Pro Ala	AA994
TTG TAC CCT CCT CCC GAA GAG CAG GTA TGT AAC CCC CTG CAG TCT CTT GAT CTT GAC AAC ACT CAG CTG GAC TTT GTG GCT ATT GTG GAT	bp3072
Leu Tyr Pro Pro Pro Glu Gly Glu Val Cys Asn Pro Leu Asp Ser Leu Asp Ser Thr Glu Leu Asp Phe Val Ala Ile Leu Asp	AA1024
GAG CCC CAG GGG CTG AGT CCT CCT CTT TCC CAT GAT CAG GGG GGC AGC TCT GGA CAT ACC CCA CCT CCT TCT GGG CCC CCA AAT GTG GCT	bp3162
Glu Pro Glu Gly Leu Ser Pro Pro Pro Ser His Asp Glu Arg Gly Ser Ser Gly His Thr Pro Pro Pro Ser Gly Pro Pro Asn Met Ala	AA1054
GTG GGC AAC ATG AGT GCT TTA CTG AGA TCC CTA CCT GGG GAA GAA TTA CTC ACT TCT AGT GCC TAA AGA GTA GGG AAT CTC ATC CAT	bp3252
Val Gly Asn Met Ser Val Leu Arg Ser Pro Pro Gly Glu Thr Glu Phe Leu Asn Ser Ser Ala ***	AA1084
CAC AGA TCG CAT TTC CTA AGG GGT TTC TAT CCT TCC AGA AAA ATT GGG GGA GCT GCA GTC CCC TGC ACA AGA TGC CCC AGG GAT GGG AGG	bp3342
TAT GGG CTG GGG GCT ATG TAT AGT CTG TAT ACG TTT TGA GGA GAA ATT TGA TAA TGA CAC TGT TTT CTG ATA ATA AAG GAA CTG CAT CAG	AA1106
AAAAAAAAAAAAAAAAAAAA	bp3432
	bp3522

Fig. 1b

used as the probe, cross-hybridizing bands were again detected in all species tested (Fig. 2c). These data suggest that the human *GLI* gene has homologues in vertebrates and invertebrates. To determine the degree of conservation, genomic clones for portions of the mouse *GLI* gene were obtained using human *GLI* cDNA sequences as hybridization probes. Sequence analysis of portions of these genomic clones revealed four regions with similarity to *GLI* that are flanked by possible splice sites¹³. These regions (*GLI* nucleotides 223–363, 535–624, 913–1077 and 1078–1308) were 84, 89, 90 and 84 percent identical at the nucleotide level and 98, 87, 98 and 92% identical at the amino

acid level, respectively (data not shown). This indicates that portions of the *GLI* gene are highly conserved between mouse and human.

The *Kruppel* gene is known to be expressed during embryogenesis and *Kruppel* mutants fail to develop all thoracic and anterior abdominal segments¹⁰. It has been shown that the *Kruppel* family mouse genes *mKrl1* and *mKrl2* are expressed in the mouse embryonal carcinoma cell line F9⁶. Therefore, it was of interest to determine whether *GLI* is expressed in similar cells. Total RNA was isolated from two human embryonal carcinoma cell lines (Tera-1 and N-Tera-2) and screened for *GLI*

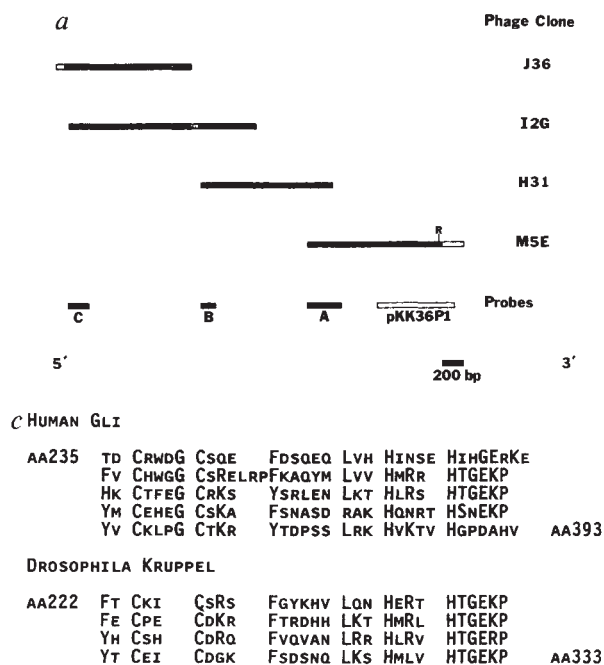


Fig. 1 Cloning and sequence analysis of *GLI* cDNAs. *a*, Map of *GLI* cDNA showing the relative position of cDNA clones and probes. Clone λM5E was isolated from the oligo(dT)-primed library and clones λH31, λI2G and λJ36 from the random hexamer primed library. Probe A was a deletion subclone of λM5E containing 326 bp of its 5' end. Probe B was an *EcoRI*/*Apal* restriction fragment of λH31 containing 150 bp from the 5' end. Probe C was a deletion subclone of λI2G containing 200 bp from the 5' end. The regions preceding and following the long open reading frame are indicated by open boxes in λJ36 and λM5E. The zinc finger region is indicated by the stippled boxes in λJ36 and λI2G. *b*, The nucleotide and amino acid sequence of the *GLI* gene derived from the four cDNA clones shown in Fig. 1*a*. *c*, The zinc finger region from human *GLI* and *Drosophila Kruppel*. Uppercase letters indicate conserved residues.

Methods. Total RNA was isolated by centrifugation through CsCl in guanidium isothiocyanate¹⁷ and polyadenylated RNA was isolated by binding to oligo(dT) cellulose. First strand cDNA was prepared using MMLV reverse transcriptase (BRL) using conditions recommended by the supplier except that random hexamer primers (Pharmacia) were used at 200 µg ml⁻¹ for the randomly primed library¹⁸. The second strand was synthesized using the method of Gubler and Hoffman¹⁹. The resulting cDNA was methylated with *EcoRI* methylase and ligated to *EcoRI* linkers (New England Biolabs). The linked cDNA was cleaved without further purification with *EcoRI* (BRL; 350 units per 2 µg of linkers), and fragments >1.0 kb were isolated following electrophoresis through a 1% agarose gel. This cDNA was ligated into alkaline-phosphatase treated, *EcoRI* digested λGT10 (Promega Biotech) and packaged with λ phage extracts (Statagene). Phage were propagated in C600 *Escherichia coli* on Luria broth agar plates. DNA from the resulting plaques was lifted with colony/plaque screen nylon membrane (NEN) and bonded by treatment with 0.4 M NaOH for 15 min. Probes for hybridization²¹ were labelled with ³²P using the random primer method²⁰. Positively hybridizing clones were purified by three rounds of plaque selection. Inserts were subcloned into the *EcoRI* site of Bluescript (Stratagene). Deletions in both directions spanning every 250 bp were generated using Exonuclease III and Mung bean nuclease²². Continuous sequence data from both strands of λM5E, λH31, λI2G and most of λJ36 were obtained from double stranded plasmid templates using a modification of the Sanger chain termination method²³. This sequence data was confirmed with single stranded templates using modified T7 DNA polymerase (Sequenase, USB)²⁴. The only discrepancy between the cDNA sequences in overlapping regions was between nucleotides 815 and 821 where λI2G had eight consecutive G residues but three other independent cDNA clones including λJ36 had only seven G residues. The insertion of this extra G in λI2G was probably the result of an error by reverse transcriptase.

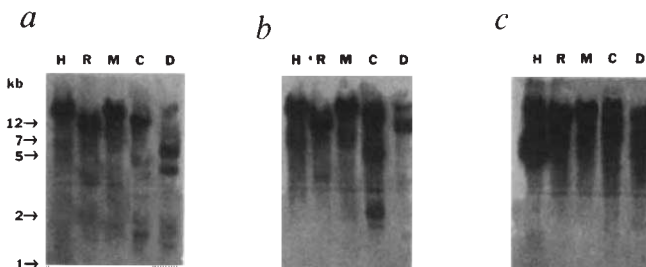


Fig. 2 Southern blot analysis demonstrating conservation of *GLI* sequences across species. DNA from normal human lymphocytes (lane H), rat spleen (lane R), mouse spleen (lane M), chicken spleen (lane C) and from the *Drosophila* Kc cell line (lane D) was probed with the *EcoRI*/*XmnI* restriction fragment (nucleotides 40 to 795) of λI2G (panel A), the *EcoRI*/*HindIII* restriction fragment (nucleotides 796 to 1189) of a deletion subclone of λI2G (panel B), or the cDNA clone λH31 (nucleotides 1177 to 2348; panel C). Vertebrate DNA was loaded at 5 µg per lane; DNA from the *Drosophila* Kc cell line was loaded at 2.5 µg per lane.

Methods. DNA was isolated, cleaved with *HindIII* (BRL), separated by electrophoresis through a 1% agarose gel and transferred to nylon membranes²⁵. Probes were labelled with ³²P by the random primer method²⁰ and hybridization was as before²¹, except that it was performed at 55 °C. Filters were given a final 60 min wash in 0.3 × SSC/0.3% SDS, 55 °C (panels *a* and *c*) or 0.3 × SSC/0.3% SDS, 55 °C (panels *a* and *c*) or 60 °C (panel *b*; 1 × SSC = 0.15 M NaCl, 15 mM sodium citrate, 1 mM Tris-HCl, pH 7.5). Molecular weights were determined from the migration of a 1 kb DNA standard ladder (BRL).

expression by RNase A protection analysis. RNA from Tera-1, N-Tera-2 and D-259 MG cells protected a 390 bp fragment (Fig. 3*a*). Comparison of autoradiograph signals from an *in vitro* transcribed *GLI* sequence suggested that Tera-1 and N-Tera-2 had five to ten *GLI* transcripts per cell. RNase A protection analysis of normal adult tissue failed to detect expression of *GLI* in brain, kidney, spleen, liver, lung, skeletal muscle, small intestinal mucosa, large intestinal mucosa, salivary glands, ovary, epididymis or placenta, but did detect low levels of expression (about one or two *GLI* transcripts per cell) in testes, myometrium, and Fallopian tubes (data not shown). The results obtained from the three cell lines by RNase A protection analysis were confirmed by Northern blot analysis. The cDNA clone λJ36 detects transcripts in Tera-1, N-Tera-2 and D-259 MG cells of ~4.0 kb in size (Fig. 3*b*). The origin of the slight difference in migration of the band in Tera-1 cells is not known but examination of the genomic organization with cDNA probes showed no obvious rearrangements.

In summary, the *GLI* gene encodes a 118Kd protein containing five repeats of the zinc finger motif and is thus distinct from other classes of genes shown to be genetically altered in cancer. By analogy to TFIIIA it seems likely that the *GLI* protein could bind DNA and potentially regulate transcription. The inclusion of *GLI* in the *Kruppel* family of zinc finger genes, as well as its conservation across species and its expression in embryonal carcinoma cells, suggest that *GLI* could have a role in normal development. The possible link between DNA binding proteins, development and oncogenesis is intriguing and is supported by several other recent observations. The proto-oncogene *c-JUN* has been reported to encode a DNA-binding protein (non-zinc finger) highly related to the transcription factor AP-1¹⁴. The proto-oncogenes *c-INT-1* and *c-REL*, although not encoding DNA binding proteins, have been shown to be homologous to the *Drosophila* segment polarity gene *Wingless*¹⁵ and the *Drosophila* maternal effect gene *Dorsal*¹⁶, respectively. Continued study of these genes should provide new insights into both neoplasia and normal development.

We thank P. White for preparing the manuscript and A. Preisinger for technical assistance. This work was supported by

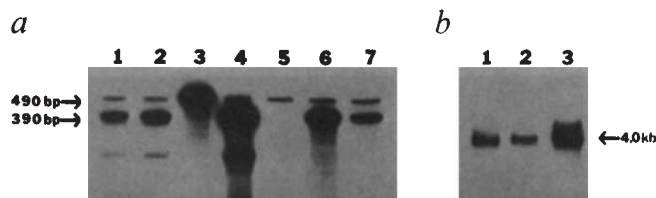


Fig. 3 Expression of *GLI* in Tera 1 and Tera 2 cells. *a*, RNase A protection analysis using a 490 bp probe that contains 390 bp of *GLI* zinc finger sequence flanked by vector. Lane 1, Tera-1 cell RNA (25 µg); lane 2, N-Tera-2 cell RNA (25 µg); lane 3, probe alone (not subjected to RNase A treatment); lane 4, D-259 MG cell RNA (9 µg); lane 5, yeast tRNA (25 µg); lane 6, 150 pg of *in vitro* transcribed 3,600 bp *GLI* transcript corresponding to ~30 copies per cell; lane 7, 30 pg of *in vitro* transcribed 3,600 bp *GLI* transcript corresponding to ~6 copies per cell. *b*, Northern blot analysis using λ J36 as probe. Lanes were loaded with poly(A⁺) RNA isolated from: Lane 1, Tera-1 cells; lane 2, N-Tera-2 cells; lane 3, D-259 MG cells.

Methods. Tera-1 cells²⁶ were from the American Type Culture Collection, N-Tera-2 cells²⁷ were provided by P. Andrews. Total RNA was isolated by the acid guanidium isothiocyanate-phenol-chloroform extraction method²⁸. The ³²P labelled RNA probe for RNase A protection was a subclone of λ I2G containing the last four zinc fingers of *GLI* (nucleotides 796-1189) cloned in Bluescript. RNase A protection was performed as described²⁹ but with the following modifications: hybridizations were in a final volume of 10 µl; only RNase A (12.5 µg ml⁻¹) was used; and the RNase A and Proteinase K incubations were at room temperature for 30 min. Calculations of copy number per cell were based on a total RNA content of 10 pg per cell. For Northern analysis poly(A⁺) RNA was isolated from 50 µg total RNA (Tera-1 and N-Tera-2) or 22 µg total RNA (D-259 MG) by selection on oligo-dT cellulose, separated by electrophoresis through a 1.5% MOPS/formaldehyde gel and transferred in dilute alkali to nylon (Bio-Rad) using conditions recommended by the supplier. The insert from a λ J36 subclone was labelled with ³²P by the random primer method²⁰ and hybridized to the filter at 50 °C in 50% formamide, 2×SSPE, 7% SDS, 0.5% nonfat dried milk, 0.2 mg ml⁻¹ salmon sperm DNA (10×SSPE = 1.8 M NaCl, 100 mM NaPO₄, 10 mM EDTA, pH 7.0). Band sizes were determined by comparison to migration of an RNA standard ladder (BRL). The previously reported size (4.8 kb, ref. 2) for the *GLI* transcript in D-259 MG cells was the result of the slower migration of transcripts in lanes containing total RNA.

National Cancer Institute training grants CA-09243 (K.W.K.) and GM-07184 (J.M.R.) and NIH grants CA-43722 (S.H.B.), NS-20023 (S.H.B.), and CA-43460(B.V.).

Received 5 January; accepted 4 February 1988.

1. Bishop, J. M. *Science* **235**, 305-311 (1987).
2. Kinzler, K. W. *et al. Science* **236**, 70-73 (1987).
3. Wong, A. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 6899-6903 (1987).
4. Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609-1614 (1985).
5. Schuh, R. *et al. Cell* **47**, 1025-1032 (1986).
6. Chowdhury, K., Deutsch, U. & Gruss, P. *Cell* **48**, 771-778 (1987).
7. Bigner, S. H., Mark, J. & Bigner, D. D. *Cancer Genet. Cytogenet.* **24**, 163-176 (1987).
8. Kozak, M. *Nucleic Acids Res.* **15**, 8125-8133 (1987).
9. Berg, J. M. *Science* **232**, 485-487 (1986).
10. Preiss, A., Rosenberg, U. B., Kienlin, A., Seifert, E. & Jackle, H. *Nature* **315**, 27-32 (1985).
11. Rosenberg, U. B. *et al. Nature* **319**, 336-339 (1986).
12. Huganir, H. L. in *Neuromodulation: The Biochemical Control of Neuronal Excitability* (eds Kaczmarek, L. K. & Levitan, I. B.) 64-85 (Oxford University Press, New York, 1987).
13. Shapiro, M. B. & Senapathy, P. *Nucleic Acids Res.* **15**, 7155-7174 (1987).
14. Bohmann, D. *et al. Science* **238**, 1386-1392 (1987).
15. Rijsewijk, F. *et al. Cell* **50**, 649-657 (1987).
16. Steward, R. *Science* **238**, 692-694 (1987).
17. Glisin, V., Crkvenjakov, R. & Byus, C. *Biochemistry* **13**, 2633-2638 (1974).
18. van Driel, I. R., Wilks, A. F., Poetersz, G. A. & Goding, J. W. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8619-8623 (1985).
19. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
20. Feinberg, A. P. & Vogelstein, B. *Anal. Biochem.* **137**, 266-267 (1984).
21. Vogelstein, B. *et al. Cancer Res.* **47**, 4806-4813 (1987).
22. Henikoff, S. *Gene* **28**, 351-359 (1984).
23. Bartlett, J. A., Gaillard, R. K. & Joklik, W. K. *Biotechniques* **4**, 210 (1986).
24. Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4767-4771 (1987).
25. Reed, K. C. & Mann, D. A. *Nucleic Acids Res.* **13**, 7207-7221 (1985).
26. Fogh, J. *Natn. Cancer Int. Monograph* **49**, 5-9 (1978).
27. Andrews, P. W. *et al. Lab. Invest.* **50**, 147-162 (1984).
28. Chomczynski, P. & Sacchi, N. *Anal. Biochem.* **162**, 156-159 (1987).
29. Winter, E., Yamamoto, F., Almoguera, C. & Perucho, M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7575-7579 (1985).

Three-dimensional NMR spectroscopy of a protein in solution

Hartmut Oschkinat*, Christian Griesinger†, Per J. Kraulis*‡, Ole W. Sørensen†, Richard R. Ernst†, Angela M. Gronenborn*‡§ & G. Marius Clore*‡§

* Max-Planck Institut für Biochemie, D-8033 Martinsried bei München, FRG

† Eidgenössische Technische Hochschule, Laboratorium für Physikalische Chemie, CH-8092 Zürich, Switzerland

The geometric information used to solve three-dimensional (3D) structures of proteins by NMR spectroscopy resides in short (<5 Å) interproton-distance data^{1,2}. To obtain these distances, the ¹H-NMR spectrum must first be assigned using correlation and nuclear Overhauser effect (NOE) experiments to demonstrate through-bond (scalar) and through-space connectivities, respectively. Because the NOE is proportional to r^{-6} , distance information can then be derived. The increased resolution afforded by extending NMR experiments into a second dimension³ enables one to detect and interpret effects that would not be possible in one dimension owing to extensive spectral overlap and much reduced information. A number of small protein structures have previously been solved in this way^{1,2}. Extending this methodology to larger proteins, however, requires yet an additional improvement in resolution as overlap of cross-peaks in the two-dimensional (2D) NMR spectra present a major barrier to their unambiguous identification. One way of increasing the resolution is to extend the 2D-NMR experiments into a third dimension. We report here the applicability of three-dimensional NMR to macromolecules using the 46-residue protein α 1-purothionin as an example.

Three-dimensional (3D) NMR experiments can be conceived by combining two 2D-NMR experiments leaving out the detection period of the first experiment and the preparation period of the second⁴. This is illustrated in a general way in Fig. 1. The location of a peak in 3D space is described by the chemical shifts of three spins, i , j and k . The information contained within a peak depends upon the nature of the experiment. For example,

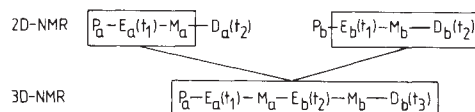


Fig. 1 Generalized representation of 2D- and 3D-NMR experiments. Abbreviations: P, preparations; E, evolution; M, mixing; D, detection.

in a 3D-NMR spectrum combining an NOE experiment with a correlation experiment, the peak at position $F1(i)$, $F2(j)$, $F3(k)$ will arise from through-space coupling between spins i and j and scalar coupling between spins j and k .

Although conceptually simple, 3D-NMR has a number of practical problems. Recording a complete 3D-NMR spectrum would be prohibitive in measurement time. For example, a 3D-NMR spectrum with a spectral width of ~8,000 Hz and a resolution of ~10 Hz in each dimension, would require ~1,000 independent increments for the first and second dimensions leading to a total minimum of 10^6 experiments. If typically a four-step phase cycle is necessary, the complete 3D-NMR experiment would take 46 days, assuming a duration of 1 s per

‡ Present addresses: Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Disorders, National Institutes of Health, Bethesda, Maryland 20892, USA (G.M.C. and A.M.G.); and Department of Molecular Biology, Biomedical Centre, Uppsala University, S-751 Uppsala, Sweden (P.J.K.).

§ To whom correspondence should be addressed.

Fig. 2 Pulse sequence used for the 3D-NOESY-HOHAHA spectrum. A minimal 32-step phase cycle was used. This included 180° phase shifts of the two π -pulses to compensate for imperfections and of the semi-selective pulses to remove axial peaks in F1 and F2, and a phase cycle selecting even coherence orders during the mixing process of the NOESY portion of the 3D experiment. A pure phase absorption spectrum was obtained using the time-proportional incrementation method³ by incrementing the phase of φ_1 and φ_2 and of φ_4 and φ_5 by 90° for every successive t_1 and t_2 value, respectively. The parameters used for the experiment shown in Fig. 2 were as follows: incremental delay (in t_1 and t_2) = $300\ \mu\text{s}$; 90° pulse (non-selective) = $22.6\ \mu\text{s}$; $90^{\text{S}1}(\text{NH}) = 2\ \text{ms}$; $90^{\text{S}2}(\text{C}^\alpha\text{H}) = 3\ \text{ms}$; SL (trim pulse) = $2\ \text{ms}$; $\tau_m(\text{NOESY}) = 200\ \text{ms}$; $\tau_m(\text{HOHAHA}) = 39\ \text{ms}$; acquisition time = $0.145\ \text{s}$. 64 independent increments were recorded for the F1 and F2 dimensions giving a total of 4,096 increments.

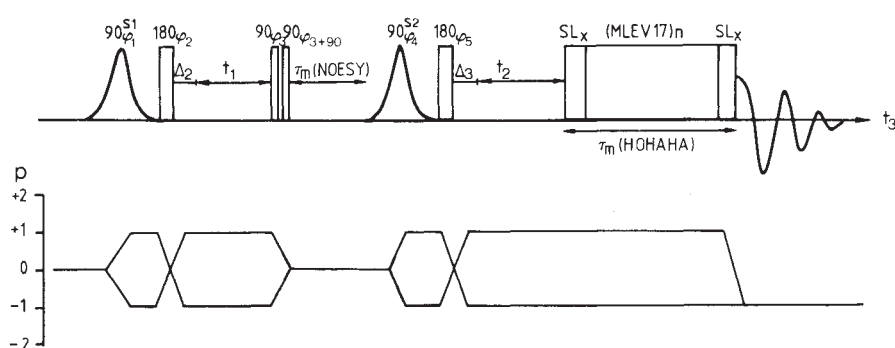
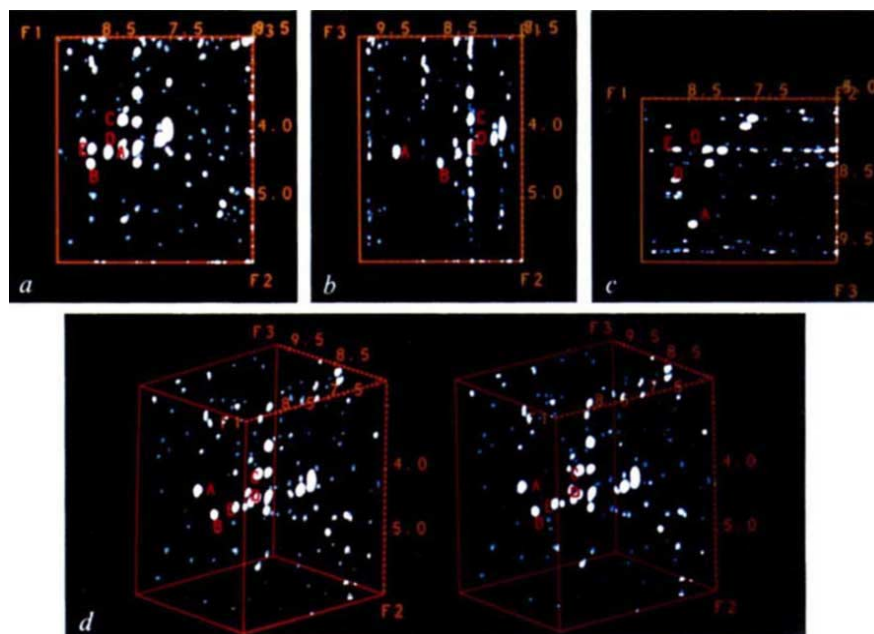
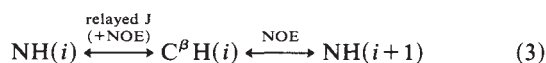
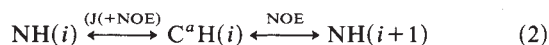
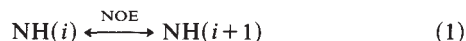


Fig. 3 NH(F1 axis)- C^αH (F2 axis)-NH(F3 axis) region of the 500 MHz 3D-NOESY-HOHAHA spectrum of 6.8 mM α 1-purothionin in 90% $\text{H}_2\text{O}/10\%$ D_2O at 25°C . The F1-F2 (a), F2-F3 (b) and F1-F3 (c) projections correspond to 2D NOESY, HOHAHA and relayed NOESY-HOHAHA spectra, respectively. d, A stereoview of the 3D spectrum. The assignments of the peaks labelled A to E are as follows: A, Ile₃₃NH(F1)-Lys₃₂ C^αH (F2)-Lys₃₂NH(F3) at 8.56, 4.31, 9.21 p.p.m.; B, Ser₃₄NH(F1)-Ile₃₃ C^αH (F2)-Ile₃₃NH(F3) at 8.80, 8.56 p.p.m.; C, Leu₃₇NH(F1)-Gly₃₆ C^αH (F2)-Gly₃₆NH(F3) at 8.32, 4.19, 8.08 p.p.m.; D, Leu₃₇NH(F1)-Gly₃₆ C^αH (F2)-Gly₃₆NH(F3) at 8.32, 3.83, 8.08 p.p.m.; E, Gln₂₂NH(F1)-Ala₂₁ C^αH (F2)-Ala₂₁NH(F3) at 8.80, 4.25, 8.10 p.p.m. (These assignments were derived from the complete assignment given in ref. 13.) The displayed spectrum, which was inspected on an Evans & Sutherland PS390 interactive graphics system using the program FRODO¹⁴, consists of $112(\text{F1}) \times 129(\text{F2}) \times 176(\text{F3})$ real data points with a digital resolution of 13.02 Hz per point in F1 and F2 and of 6.88 Hz per point in F3. The experimental conditions were: 6.8 mM α 1-purothionin in 90% $\text{H}_2\text{O}/10\%$ D_2O containing 500 mM sodium phosphate buffer pH 4.0. The spectrum was recorded on a Bruker AM500 spectrometer equipped with a selective excitation unit supplemented by a 30 W linear amplifier.



scan. This problem can be eliminated by recording only part of the 3D spectrum (that is, a sub-volume of the total spectrum) using semi-selective excitation pulses⁴⁻⁶.

For the purposes of illustration, we present a sub-spectrum comprising the NH- C^αH region to demonstrate the utility of 3D-NMR in sequential resonance assignment. The general strategy for sequential assignment involves the identification of the following through-space (NOE) and scalar (J) pathways:



Through-space connectivities are detected by the NOESY (two-dimensional nuclear Overhauser spectroscopy) experiment^{7,8}. Through-bond connectivities can be measured by correlated spectroscopy (COSY⁹) that delivers exclusively direct connectivities. A more versatile approach, however, is total-correlated spectroscopy, known as TOCSY¹⁰ or homonuclear Hart-

mann-Hahn (HOHAHA)¹¹ spectroscopy, which detects successively direct, single and multiple relayed scalar connectivities as the mixing time is increased. The 3D-NMR experiment selected for illustration was a NOESY-HOHAHA with detection of inter- and intra-residue NOE connectivities between NH and C^αH protons in the first part of the experiment and J-coupling intra-residue connectivities between $\text{C}^\alpha\text{H}(i)$ and $\text{NH}(i)$ protons in the second part. This is achieved using semi-selective Gaussian excitation pulses¹² centred in the middle of the NH and C^αH regions for the preparation pulse and the last pulse of the mixing process, respectively, in the NOESY part of the sequence. In this way a 3D spectrum is obtained with restricted spectral widths in the F1 and F2 dimensions and the full spectral width in the F3 dimension. The experimental scheme is shown in Fig. 2.

The 3D-NOESY-HOHAHA experiment was recorded in ~69 h at 500 MHz on a 6.8 mM sample of α 1-purothionin in 90% $\text{H}_2\text{O}/10\%$ D_2O at 25°C . Recording the spectrum in H_2O is required to observe the exchangeable NH protons. The suppression of the huge water signal (~100 M) renders the experiment technically demanding. This was achieved by on-resonance irradiation during the relaxation delay and during the NOESY

mixing period. The NOESY and HOHAHA mixing times were 200 ms and 39 ms, respectively.

The results of this experiment are illustrated in Fig. 3. The F1-F2 projection (Fig. 3a) corresponds to a NOESY spectrum of the NH(F1 axis)-C α H(F2 axis) region; the F2-F3 projection (Fig. 3b) to a HOHAHA spectrum of the C α H(F2 axis)-NH(F3 axis) region; and finally the F1-F3 projection (Fig. 3c) to a relayed NOESY-HOHAHA spectrum of the NH(F1 axis)-NH(F2 axis) region. A stereoview of the cube is shown in Fig. 3d. Two sorts of peaks are present: auto-peaks which arise solely from intra-residue effects, and combined intra- and inter-residue-effect peaks. The former have the chemical shift of the NH(*i*) proton in the F1 and F3 dimensions and the chemical shift of the C α H(*i*) proton in the F2 dimension, whereas the latter have the chemical shifts of the NH(*i*+1), C α H(*i*) and NH(*i*) protons in the F1, F2 and F3 dimensions, respectively. The assignments of five such combined intra- and inter-residue peaks are given in Fig. 3. It will be noted that the strongest peaks of that kind arise from β -strand structures. This is because the contribution from the scalar C α H(*i*)-NH(*i*) and NOESY C α H(*i*)-NH(*i*+1) interactions are very large in such cases owing to sizeable values for the coupling constant $^3J_{\text{HN}\alpha}$ and very short C α H(*i*)-NH(*i*+1) distances (~ 2.2 Å), respectively.

Informative 3D-NMR spectra of proteins can be obtained in a reasonable time with concentrations of material no higher than those regularly used in 2D-NMR experiments. What do such 3D-NMR experiments offer? In the case presented here they provided separation of C α H(*i*)-NH(*i*+1) inter-residue NOEs in a third dimension according to the chemical shift of the NH proton of residue *i*, thereby permitting us to identify a number of inter-residue effects that could not be assigned by 2D-NMR owing to overlap with intra-residue C α H(*i*)-NH(*i*) cross-peaks. For example, in the 2D-NOESY experiments, the C α H(*i*)-NH(*i*+1) cross peaks between Ala 21 and Gln 22 (4.25, 8.80 p.p.m.), Ile 33 and Ser 34 (4.46, 8.80 p.p.m.), and Gly 36 (C α H) and Leu 37 (4.19, 8.32 p.p.m.) are superimposed on the intra-residue C α H(*i*)-NH(*i*) cross-peaks of Asn 14 (4.22, 8.80 p.p.m.), Ser 35 (4.48, 8.81 p.p.m.) and Leu 37 (4.20, 8.32 p.p.m.), respectively¹³. In the 3D-NOESY-HOHAHA spectrum, however, the chemical shifts of these two sets of peaks differ in the F3 dimension. Thus the inter-residue cross-peaks in the 2D-NOESY spectrum give rise to mixed inter- and intra-residue cross-peaks in the 3D-NOESY-HOHAHA spectrum (Fig. 3) with NH proton chemical shifts of Ala 21 (8.10 p.p.m.; peak E), Ile 33 (8.56 p.p.m.; peak B) and Gly 36 (8.08 p.p.m.; peak D), respectively, in the F3 dimension, which differ from those of the auto-peaks of Asn 14 (8.80 p.p.m.), Ser 35 (8.81 p.p.m.) and Leu 37 (8.32 p.p.m.), respectively.

Many further useful 3D-NMR experiments can be envisaged. A particularly useful, although technically difficult, 3D-NMR experiment seems to be a HOHAHA-NOESY performed in H₂O with C α H and NH protons in the F1 and F2 dimensions, respectively, which would provide the crucial NH-aliphatic proton NOEs. Other potentially useful experiments would include the combination of the NOESY experiment with a heteronuclear correlation experiment; two such experiments would be to spread the NH-aliphatic and C α -NH/aliphatic NOEs in a third dimension according to the ¹⁵N and ¹³C chemical shift, respectively, of the directly bonded heavy atom.

This work was supported by the Max-Planck Gesellschaft and the Bundesministerium für Forschung und Technologie (G.M.C. and A.M.G.). P.J.K. is supported by the Swedish Natural Science Research Council. We thank Alwyn Jones for helpful discussions concerning the displaying of 3D-NMR spectra.

Received 14 December 1987; accepted 3 February 1988.

- Griesinger, C., Sørensen, O. W. & Ernst, R. R. *J. Am. chem. Soc.* **109**, 7227-7228 (1987).
- Braunschweiler, R., Madsen, J. C., Griesinger, C., Sørensen, O. W. & Ernst, R. R. *J. magn. Reson.* **73**, 380-385 (1987).
- Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. *J. chem. Phys.* **71**, 4546-4553 (1979).
- Anil Kumar, Ernst, R. R. & Wüthrich, K. *Biochem. biophys. Res. Commun.* **95**, 1-6 (1980).
- Aue, W. P., Bartholdi, E. & Ernst, R. R. *J. chem. Phys.* **64**, 2229-2246 (1976).
- Braunschweiler, R. & Ernst, R. R. *J. magn. Reson.* **53**, 521-528 (1983).
- Davis, D. G. & Bax, A. *J. Am. chem. Soc.* **107**, 2821-2822 (1985).
- Bauer, C., Freeman, R., Frenkiel, T., Keeler, J. & Shaka, A. J. *J. magn. Reson.* **58**, 442-457 (1984).
- Clare, G. M. *et al. J. molec. Biol.* **193**, 571-578 (1987).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268-272 (1978).

Heparan sulphate bound growth factors: a mechanism for stromal cell mediated haemopoiesis

R. Roberts, J. Gallagher*, E. Spooncer, T. D. Allen, F. Bloomfield & T. M. Dexter

Departments of Experimental Haematology and *Medical Oncology, Paterson Institute for Cancer Research, Christie Hospital and Holt Radium Institute, Withington, Manchester M20 9BX, UK

The proliferation and development of haemopoietic stem cells takes place in close association with marrow stromal cells^{1,2}. This intimate cell contact presumably enables the stem cells and their progeny to respond to stimuli present on the stromal cell surface. While the nature of these stimuli has not been determined, it is likely that growth factors play some role. Recently, it was demonstrated that the natural and the recombinant haemopoietic growth factor, granulocyte/macrophage colony stimulating factor (GM-CSF), could be adsorbed out of solution by an extract of human marrow stromal extracellular matrix (ECM) with retention of biological activity³. However, the precise ECM molecules involved were not identified. Here, we clearly demonstrate that the major sulphated glycosaminoglycan of mouse marrow stroma, heparan sulphate⁴, possesses the ability to adsorb both GM-CSF and the multilineage haemopoietic growth factor, Interleukin 3 (IL-3). Furthermore, these growth factors, once bound, can be presented in the biologically active form to haemopoietic cells.

In these experiments, we used two growth-factor-dependent multipotent cell lines: FDCP-Mix A4 which responds to IL-3 and FDCP-Mix A7 which responds to both to IL-3 and GM-CSF⁵. To examine the role that glycosylation of growth factors plays in the ability to bind to components of the ECM, both native 'n' (glycosylated) and *Escherichia coli*-derived recombinant 'r' (non-glycosylated) IL-3 and GM-CSF were used. Initially, we examined the ability of a commercially available ECM preparation (Matrigel) to bind and present these growth factors. The data (Fig. 1a and b) clearly show that when FDCP-Mix A4 and FDCP-Mix A7 are added to Matrigel which has been pre-incubated with IL-3 or GM-CSF, they can survive and proliferate.

We could exclude the possibility that cell survival occurred in response to growth factor released by diffusion from Matrigel because the overlying medium from growth-factor-treated Matrigel plates was unable to support the growth of IL-3 or GM-CSF-dependent cells. We conclude, therefore, that the stimulus for growth is *bound* growth factor. Significantly, both native and recombinant growth factor could support the growth of the stem cells, indicating that the protein portion of the growth factor is responsible for binding. It is also clear that the bound growth factors continue to express functional domains for interaction with receptors on the target cells.

We investigated whether enzymatic digestion of the glycosaminoglycan components of the Matrigel had an effect on the ability of the ECM to bind and present growth factors. Two enzymes with differing specificities were used: heparitinase, which digests heparan sulphate and chondroitinase ABC, which digests chondroitin sulphate. The data (Fig. 2a and b) show

1. Wüthrich, K. *NMR of Proteins* (Wiley, New York, 1986).

2. Clare, G. M. & Gronenborn, A. M. *Protein Engineering* **1**, 275-288 (1987).

3. Ernst, R. R., Bodenhausen, G. & Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Clarendon, Oxford, 1987).

4. Griesinger, C., Sørensen, O. W. & Ernst, R. R. *J. magn. Reson.* **73**, 574-579 (1987).

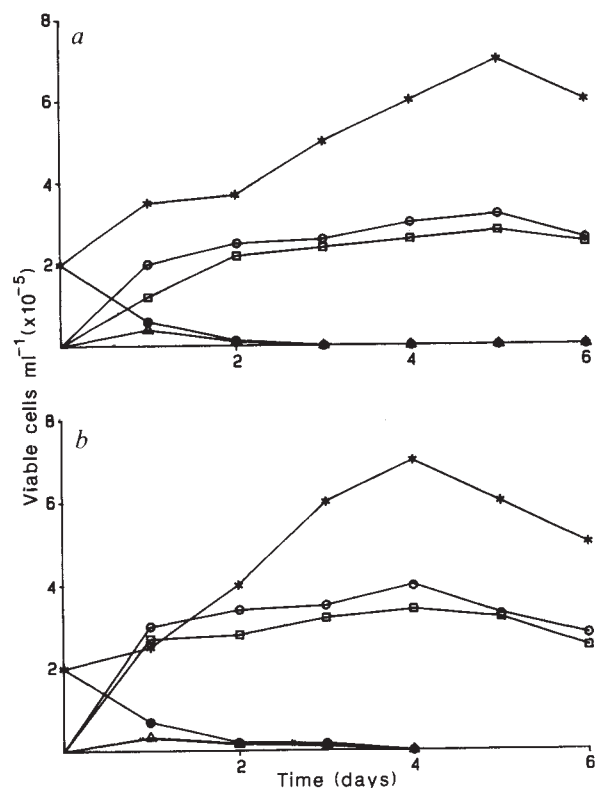


Fig. 1 Cell proliferation on growth-factor-treated Matrigel. *a*, FDCP-Mix cells. $\square$, nIL3-treated Matrigel; $\circ$, rIL3-treated Matrigel; $\triangle$, untreated Matrigel; *, positive control (liquid culture in the presence of nIL3); $\bullet$, negative control (liquid culture in the absence of nIL3). *b*, FDCP-Mix A₇ cells; $\square$, nGM-CSF-treated Matrigel; $\circ$, rGM-CSF-treated Matrigel; $\triangle$, untreated Matrigel; *, positive control (liquid culture in the presence of nGM-CSF); $\bullet$, negative control (liquid culture in the absence of nGM-CSF).

Methods. 35-mm diameter plastic Petri dishes were coated with 0.5 ml of the basement membrane extract Matrigel (ref. 6, Collaborative Research Ltd.) warmed to 37 °C for 30 min to permit matrix formation, then washed four times in Fischer's medium supplemented with 20% (v/v) horse serum (FHS). These were then treated with growth factors by incubating either 1 ml of medium conditioned by the cell line WEHI (a source of nIL-3), 1 ml of lung-conditioned medium (a source of nGM-CSF), 1 ml of rIL3 (10^{-3} U ml $^{-1}$ in PBS, Biogen), 1 ml of rGM-CSF (50 U ml $^{-1}$ in PBS, Biogen) or 1 ml of FHS (control) over the matrices for 3 h, at 37 °C washing them four times in FHS and inoculating them with 2×10^5 washed FDCP-mix or FDCP-Mix A₇ cells in 1 ml FHS. Growth medium was removed after 2 h and the adherent cells (30–50% of the initial population) were overlaid with fresh growth medium. Aliquots of medium were removed at specific times and the non-adherent cells counted. About 50% of the cells remained attached to the Matrigel during the time course of the experiment. Three separate experiments were performed, each in duplicate. Standard errors were $\pm 5\%$.

that digestion with heparitinase abolishes the ability of Matrigel to bind and/or to present growth factor to cells, whereas treatment with chondroitinase ABC has no effect. From these data, we conclude the heparan sulphate is intimately involved in the binding/presentation process.

Because Matrigel is a complex ECM extract containing collagen, fibronectin and proteoglycans⁶, it was possible that binding and presentation of growth factor by the ECM requires more than one component of this ECM. Also the growth-factor-binding properties of Matrigel may not be present in the extracellular components of the stromal cells that are able to support haemopoiesis. To address this question, we isolated the heparan sulphate chains from proteoglycans produced by the primitive mesenchymal cell line, 3T3, which we have shown can support

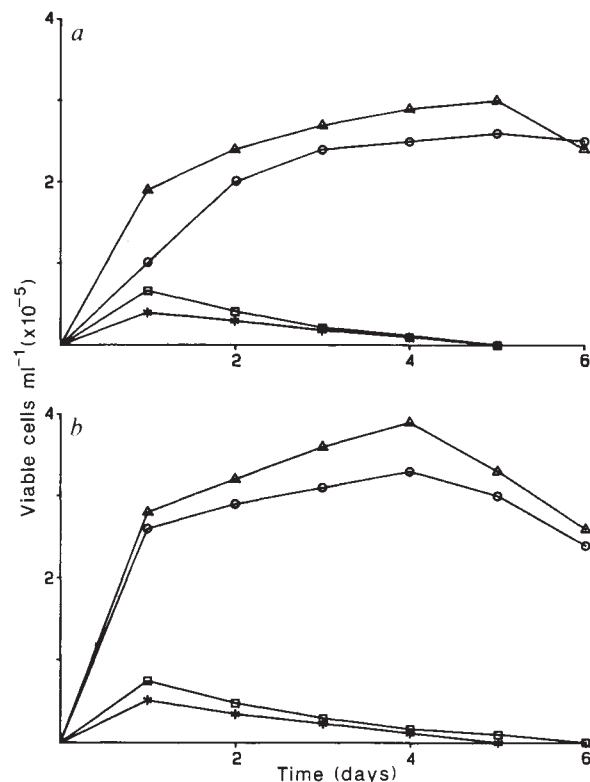


Fig. 2 Cell proliferation on Matrigel matrices digested with heparitinase (Hep) or chondroitinase ABC (Chon) prior to growth factor treatment. *a*, FDCP-Mix cells. *, Hep digested, nIL3-treated Matrigel; $\square$, Hep digested, rIL3-treated Matrigel; $\circ$, Chon digested, nIL3-treated Matrigel; $\triangle$, Chon digested, rIL3-treated Matrigel. *b*, FDCP-Mix A₇ cells. *, Hep digested, nGM-CSF-treated Matrigel; $\square$, Hep digested, rGM-CSF-treated Matrigel; $\circ$, Chon digested, nGM-CSF-treated Matrigel; $\triangle$, Chon digested, rGM-CSF-treated Matrigel. Matrigel matrices (Fig. 1) were washed with two changes of Hep buffer (0.25 M Na acetate, 0.25 mM Ca acetate, 100 μ g ml $^{-1}$ BSA, pH 7.0) or Chon buffer (50 mM TrisHCl, 50 mM NaCl, 100 μ g ml $^{-1}$ BSA, pH 7.8) as appropriate. Hep (ICN Biomedical) was added at 5 mU ml $^{-1}$ in Hep buffer or Chon (Miles Chemicals Ltd) was added at 0.25 U ml $^{-1}$ in Chon buffer. The digestions were carried out for 4 h at 43 °C and 37 °C respectively. After digestion the matrices were washed in FHS, treated with growth factors and inoculated with cells (Fig. 1).

the proliferation and development of both freshly isolated stem cells and the multipotent growth-factor-dependent cell lines used in this study⁷. To immobilize this heparan sulphate in the absence of collagen and fibronectin (and thus rule out the possibility that binding of growth factor requires co-operative interactions between ECM components), we covalently linked the 3T3 heparan sulphate to the lipid L- α -phosphatidyl ethanolamine dipalmitoyl (PPEADP)⁸. This heparan sulphate/lipid complex readily adhered to plastic culture dishes. Figure 3*a* and *b* shows that when this complex was pre-incubated with either native or recombinant growth factor, it could support the growth of the factor-dependent stem cells. This ability was lost if the complex was treated with heparitinase but was unaffected by chondroitinase ABC. Thus, isolated heparan sulphate chains can bind and present both native and recombinant haemopoietic growth factors in the active form and other extracellular components are not required.

We have previously shown that stem cell growth and the development of stromal cell monolayers requires direct cell to cell contact; no growth is seen in intercellular regions of the stromal cell cultures⁷. In these regions, however, there are extensive deposits of ECM^{9,10}. To determine whether the intercellular ECM is functionally different from the stromal cell membrane,

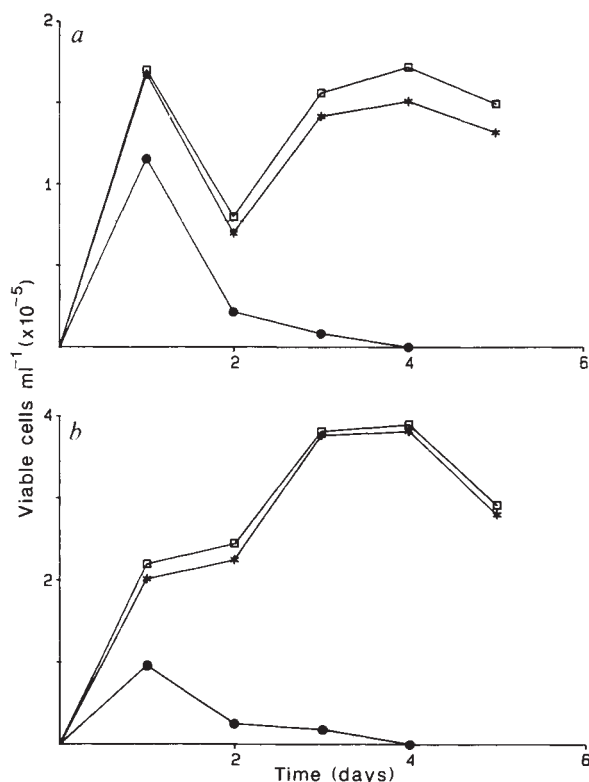


Fig. 3 Effect of glycosaminoglycan degrading enzymes on the ability of PPEADP-heparan sulphate complexes to bind and present growth factors. *a*, FDCP-Mix cells, *b*, FDCP-Mix A₇ cells. Enzyme treatment: ●, Hep; *, Chon; □, enzyme buffer only (i.e. undigested controls).

Methods. PPEADP-heparan sulphate complexes were adsorbed onto 35 mM plastic Petri dishes. The complexes were incubated with specific glycosaminoglycan-degrading enzymes and then treated with rIL3 for growth of FDCP-Mix cells or rGM-CSF for growth of FDCP-Mix A₇ cells (Fig. 1). The dishes were inoculated with 2×10^5 cells in 1 ml FHS and viable, non-adherent cells were determined at 24 h intervals. Data are for recombinant growth factors but results were similar with natural counterparts. The heparan sulphate proteoglycans were isolated from 3T3 cell cultures¹³. Heparan sulphate chains were released from protein by treatment with 50 mM NaOH and covalently linked to PPEADP⁸.

we prepared cell-free layers of both 3T3 and marrow-stromal ECM¹¹. The ECM was then pre-incubated with IL-3 or GM-CSF and subsequently tested for its capacity to support the growth of IL-3 or GM-CSF dependent cell lines. No growth was seen (data not shown), indicating that intercellular and non-membrane-associated ECM cannot bind and/or present active growth factor. This presumably explains the importance of direct contact between haemopoietic cells and the stromal cell surface.

Since haemopoiesis occurs in stromal cell co-cultures in the absence of added growth factors such as IL-3 or GM-CSF, it appears that the growth factors necessary for haemopoiesis are being synthesized by the stromal cells and are then transferred to the stromal cell plasma membrane. The growth factors are not released into the culture medium⁷ and our findings suggest they may be retained on the membrane by binding to membrane-associated heparan sulphate. This, in turn, suggests a more general mechanism by which certain growth factors (particularly those that are not normally found in the circulation) are stored by the ECM, when their activity is localized and perhaps regulated by ECM components such as heparan sulphate¹⁴. Apart from the haemopoietic cell growth factor discussed here, basic fibroblast growth factor (bFGF), a potent endothelial cell mitogen, has recently been identified in the basement membrane of cornea and vascular endothelial cells, where it may be bound to heparan sulphate¹⁵. It is also possible that binding to heparan

sulphate can protect growth factors such as GM-CSF and IL-3 from enzymatic degradation, as has been found for FGF¹⁶. In this context, it is perhaps significant that haemopoietic stem cells do not possess their own cell surface heparan sulphate although they do produce a chondroitin-4-sulphate proteoglycan (Gallagher *et al.*, in preparation). This molecule, however, was unable to bind growth factors (data not shown), ruling out the possibility that stem cells can bind growth factors via their surface polysaccharides. Taken together, these data provide a mechanistic explanation for the observed total dependence of haemopoietic cells on intimate stromal cell contact. While the relationship of stromal cell mediators of growth to the defined growth factors such as IL-3 and GM-CSF is obscure¹², the present work confirms and extends previous studies³ in providing an explanation of the events observed in stromal cell co-culture systems. Furthermore, since haemopoiesis *in vivo* also occurs in association with stromal cells, the results are likely to be of physiological significance.

Supported by the Cancer Research Campaign (UK).

Received 14 December 1987; accepted 11 February 1988.

- Dexter, T. M., Spooner, E., Schofield, R., Lord, B. I. & Simmons, P. *Blood Cells* **10**, 315-339 (1984).
- Bentley, S. A. *Exp. Haematol.* **9**, 308-312 (1981).
- Gordon, M. Y., Riley, G. P., Watt, S. M. & Greaves, M. F. *Nature* **326**, 403-405 (1987).
- Gallagher, J. T., Spooner, E. S. & Dexter, T. M. *J. cell Sci.* **63**, 155-171 (1983).
- Spooner, E., Heyworth, C. M., Dunn, A. & Dexter, T. M. *Differentiation* 111-118 (1986).
- Kleinman, H. K. *et al. Biochemistry* **132**, 312-318 (1986).
- Roberts, R. A. *et al. J. cell Physiol.* **132**, 203-214 (1987).
- Tang, P. W., Gooi, H. C., Hardy, M., Lee, V. C. & Feizi, T. *Biochem. Res. Commun.* **132**, 474-480 (1985).
- Dexter, T. M., Simmons, P., Purnell, R. A., Spooner, E. & Schofield, R. *Aplastic Anaemia; Stem Cell Biology and Advances in Treatment* (eds Young, N. S., Levine, A. & Humphries, R. K.) 13-33 (Liss, New York, 1984).
- Zuckerman, K. S. & Wicha, M. S. *Blood* **61**, 540-547 (1983).
- Todaro, G. & Green, H. *J. cell Biol.* **17**, 229-313 (1983).
- Dexter, T. M. *J. cell Science* **88**, 1-6 (1987).
- Gallagher, J. T. & Walker, A. *Biochem. J.* **230**, 665-674 (1985).
- Gallagher, J. T., Lyon, M. & Steward, W. P. *Biochem. J.* **236**, 313-325 (1986).
- Vlodavsky, I., Sullivan, R., Sasse, J., Folkman, J. & Klagsbrun, M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2292-2296 (1987).
- Gospodarowicz, D. & Cheng, J. *J. cell Physiol.* **128**, 475-484 (1986).

Receptor-directed focusing of lymphokine release by helper T cells

Wen-Jen Poo, Lisa Conrad & Charles A. Janeway Jr

Department of Pathology and Section of Immunobiology, Howard Hughes Medical Institute at Yale University School of Medicine, New Haven, Connecticut 06510, USA

The interaction between helper T cells and B cells, leading to the production of antibody to thymus-dependent antigens, was the first cell interaction clearly defined in the immune system¹⁻³; it remains both paradigmatic and controversial. Two requirements of this interaction, that the helper cell (T_H) and the B cell must recognize antigenic determinants that are physically linked^{4,5}, and that the T_H and the B cell must share genes encoding major histocompatibility complex (MHC) class II molecules⁶, led to the concept that T_H-B interaction required an intimate physical association of the two cell types. But *in vitro* studies have shown that T_H can be replaced by soluble, antigen-nonspecific factors, capable of activating any B cell to secrete antibody^{7,8}. We have previously proposed that the requirements for T_H-B contact might result from T_H cells releasing their lymphokines in a polar fashion directed at that portion of the cell membrane where T-cell receptor cross-linking is actually occurring⁹⁻¹⁶. Using an artificial monolayer of a cloned helper T-cell line, we show that lymphokines are released preferentially over the area of receptor cross-linking under conditions of limited T_H-cell activation. Thus, it appears that one important aspect of the specificity of T_H-B cell interactions is the receptor-directed polar release of helper lymphokines.

We have previously described a cloned helper T-cell line, termed D10.G4.1 (D10) derived from AKR/J mice and specific

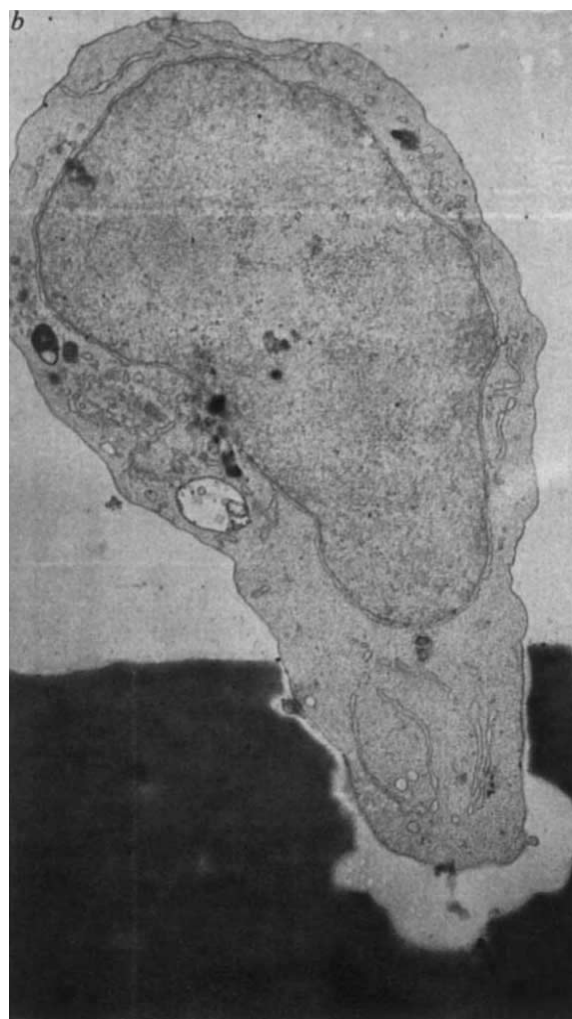
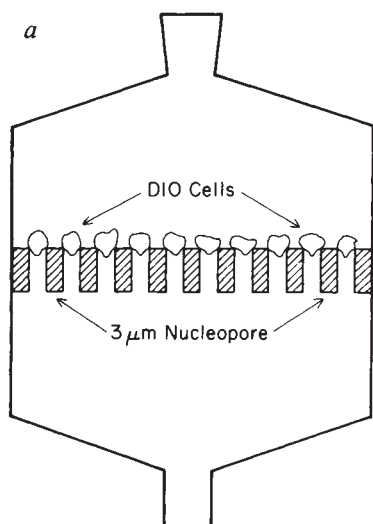


Fig. 1 The experimental system used in these studies. *a*, Schematic of the culture system. *b*, Electron micrograph of a D10 cell in the pore on the top of a nucleopore filter. Filters were fixed in glutaraldehyde, embedded in EPON and sectioned. No cells were observed on the bottom face of the filter or associated with the surface away from pores.

Methods. Swinnex filter holders (Millipore) 13 mm in diameter were mounted with Nucleopore chemotaxis membranes having a 3- μ m pore size. Suspensions of D10 cells at 10^6 ml $^{-1}$ in balanced salt solution were centrifuged at 600g onto the membranes until fluid would not pass under positive pressure. The filter chambers top and bottom were then washed using fine flexible tubing and a syringe to remove cells not firmly lodged in the openings of the filter. Both sides of the chamber were then loaded with Click's EHAA medium (Gibco) containing 10% fetal calf serum, and monoclonal anti-D10 receptor antibody 3D3 was added at various concentrations to one side of the chamber. Alkaline phosphatase was added at 5 μ g ml $^{-1}$ to the same side of the chamber as 3D3. Both ends were then sealed and the chambers cultured upright at 37 °C for 4–16 h. The seals were then removed and culture fluid collected separately from both sides of the membrane. This was analysed for alkaline phosphatase by development of the substrate *p*-nitrophenyl phosphate, disodium salt (Sigma). Chambers with more than 15% diffusion of alkaline phosphatase were discarded from consideration. The supernatants from both halves of the chamber were then assayed for release of T-cell growth factor (TCGF, in this case IL-4) by titration on HT-2 cells at 10^4 per well as previously described.

for the antigen conalbumin (CA; ref. 17) in association with I-A^k (refs 18–20). We have prepared a monoclonal antibody (3D3) specific for the receptor on this cloned T-cell line¹⁸. Unlike most anti-T-cell receptor antibodies, this antibody will activate the cloned D10 cell line at very low concentrations in the soluble, bivalent form, although the monovalent Fab fragment is not stimulatory¹⁹. This is a property of the D10 cell line rather than the 3D3 antibody, as monoclonal anti-T-cell receptor antibodies directed at shared T-cell receptor determinants stimulate D10 but inhibit other cloned lines in our hands and those of others²¹. Furthermore, 3D3 triggers potent lymphokine release by D10 cells²⁰, and D10-dependent activation of B cells^{9,18}. The monoclonal antibody precipitates the expected disulphide-linked heterodimeric molecule from D10 cells¹⁹. Finally, in morphological experiments, we have shown that CA:I A^k-bearing cells conjugate with D10 cells, leading to a polarization of the cytoskeleton²² and of the T-cell receptor-CD4 complex^{23,24} of the D10 cells within 15 minutes. Thus, this cell line appeared to be an ideal model for analysing the existence of polar release of lymphokines by helper T cells.

To examine this question, we devised a culture system illustrated schematically in Fig. 1*a*, and shown microscopically in Fig. 1*b*. In this system, D10 cells are washed, suspended in medium, and forced into the pores of 3- μ m nucleopore filters

housed in sterile Swinnex chambers. When the membrane becomes hydrostatically sealed by D10 cells, both sides of the chamber are carefully washed with medium, and the medium on one side is replaced with medium containing 3D3 and alkaline phosphatase as a biologically neutral and readily detected diffusion marker. After 6–24 h, medium is collected from both sides of the chambers and alkaline phosphatase is measured in both supernatants. In those chambers in which the leakage of alkaline phosphatase is less than 15%, we then assay for lymphokine release by the ability of the supernatants to promote the growth of the HT-2 cell line. This assay measures T-cell growth factor production; in the case of D10, this T cell growth factor is IL-4 as shown by inhibition with monoclonal anti-IL-4 antibody^{25,26}. Thus, we are measuring a known B-cell helper factor^{25–27}.

As can be seen in Fig. 2*a*, lymphokine release is greater in the top chamber than it is in the bottom chamber in experiments using 100 ng ml $^{-1}$ 3D3 to stimulate D10, whether the antibody is added to the top or to the bottom chamber; stimulation of D10 is clearly more efficient when the 3D3 is added to the top chamber. This shows that diffusion of lymphokine is limited by the membrane occluded by the cells. The excess of lymphokine in the top chamber in all cases probably reflects the much greater proportion of the D10 cell body above the membrane, as seen in Fig. 1*b*. But the ratio of lymphokine in the top to that in the bottom chamber is the same, regardless of where the antibody is added. Thus, at this dose of antibody, which represents about a 100-fold excess¹⁹, no polarity in lymphokine release is observed. But when the dose of antibody is reduced to 20 ng ml $^{-1}$,

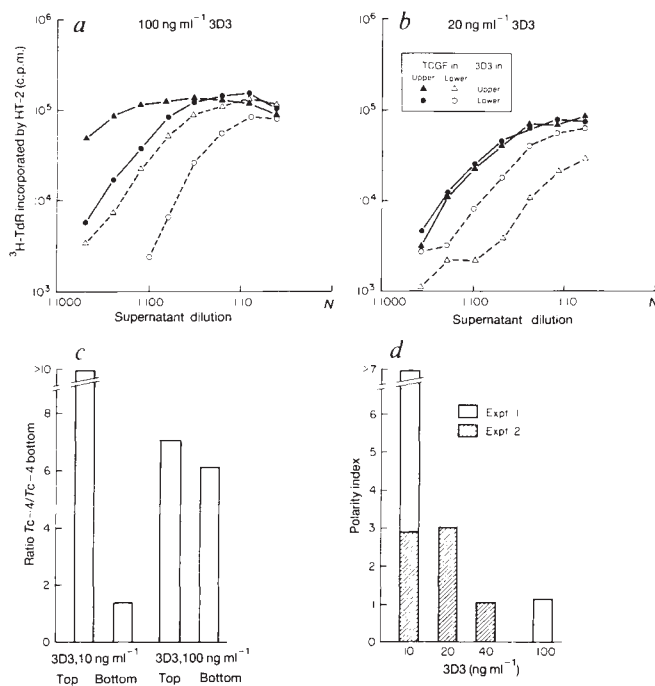


Fig. 2 Analysis of IL-4 release by D10 cells stimulated in the two-chamber system described in Fig. 1. *a*, Titration on HT-2 cells of IL-4 activity from D10 cells stimulated by 100 ng ml⁻¹ 3D3 in a 16-h culture. *b*, As for *a*, but using 20 ng ml⁻¹ 3D3 in a 20-h culture. *c*, The ratio of IL-4 in the upper and lower chambers as a function of 3D3 location and dose. Ratio is calculated from TCGF levels in which units are calculated as the reciprocal of that dilution of IL-4 that gives half-maximal activation of HT-2 cells as determined from *a*, *b*, and similar titration curves. *d*, Effect of 3D3 dose on the polarity of lymphokine release by D10 cells. Two experiments are shown. Polarity index is calculated from the ratio of top to bottom lymphokine release shown in *c* using the formula: polarity index = (IL-4 top/IL-4 bottom ratio 3D3 in top chamber)/(IL-4 top/IL-4 bottom ratio, 3D3 in bottom chamber). As such, the polarity index corrects for the differential distribution of cell body on the two sides of the membrane dividing the chamber. Indices greater than 1 indicate polar release of lymphokine.

leading to a reduction in net lymphokine secretion, we now observe strong polarity in lymphokine release (Fig. 2c). Thus, polarity does occur, but not when the stimulation is maximal. This experiment has been repeated at several doses of 3D3, as shown in Fig. 2d, with similar results. Control experiments examining diffusion of IL-4 as D10 supernatant added directly to the top chamber in the absence of cells showed essentially complete equilibration within 16 h. Thus, the effects observed cannot reflect differential absorption of IL-4 to one or another face of the symmetrical nucleopore filter (data not shown). Also, cells clearly limit diffusion in our system. Finally, no lymphokine was produced when supernatants were centrifuged, and the pellets, containing any cells that might migrate off the filter, were cultured for a further 24 h.

The finding that polarity of lymphokine release could be demonstrated in this system, and that polar lymphokine release resulted from T-cell receptor cross-linking, is consistent with our previous finding that D10 cells polarize their cytoskeleton upon encountering B lymphoma cells pulsed with specific antigen and bearing the appropriate MHC molecules^{22,23}. It seemed surprising, however, that increasing the level of stimulation of D10 cells led to a loss of polarity. Although this result could be accounted for by a minor leak of 3D3 across the membrane, the symmetry of lymphokine release under conditions of maximal stimulation could also be due to a loss of polarity under such conditions. Reducing the duration of incubation with 3D3 did not affect the data, indicating that diffusion of

3D3 is unlikely to account for this effect. Rather, under conditions of supra-optimal T-cell activation, polarity of lymphokine release may be lost, leading to nonspecificity in cell interactions.

These studies go some way towards explaining how T_H cells can use nonspecific mediator molecules to activate specific B cells. We would propose that the T_H receptor recognizes antigen-MHC complexes on the specific B cell surface, and that the T_H cell thereby becomes activated to release helper lymphokines. Specific B-cell capture of antigen via the immunoglobulin receptor is clearly important in the specificity of this interaction^{28,29}. When the antigen-MHC stimulus engages only a fraction of the receptors of the T_H cell, as is virtually certain to occur *in vivo* or in *in vitro* cultures carried out at limiting antigen concentration, then the lymphokines will be released focally in the area of T cell-B cell contact, due to activation via the cross-linking of the T-cell receptor in this region of the T_H cell membrane. Consistent with this, no free lymphokine is detected during physiological B-cell activation by D10 T_H cells (J. P. Tite, K. Bottomly and C.A.J., unpublished results).

It should be noted that these studies address IL-4 release only; further analyses will be necessary to confirm whether other lymphokines are also released in a polar fashion. IL-4 release by D10 cells is detectable within 60–90 min of exposure to 3D3; thus, for polar release of IL-4 to be relevant, specific T_H-B interaction must be sustained for a similar period of time. The detailed cellular mechanisms by which this focal release of lymphokine occurs need to be explored, and the experimental system presented here can be used to do so. The polar release of nonspecific mediator molecules provides yet another analogy between the immune system and the nervous system, and reinforces the impression that one should consider the lymphokines involved in specific cell interactions as transmitters rather than as hormones.

This work was supported in part by the NIH and by the Howard Hughes Medical Institute. C.A.J. is an investigator of the Howard Hughes Medical Institute, L.C. is a summer Yale medical student, and W.-J.P. is supported by an NIH training grant. We thank Abraham Kupfer, S. John Singer, Zaid Smith, James Jamieson and John Tite for helpful discussions, Jonathan Kaye for the D10 cells and 3D3 antibody, Patricia Conrad and Jo Ann Buchanan for technical assistance and Liza Cluggish for typing the manuscript.

Received 30 December 1987; accepted 2 February 1988.

- Claman, H. N., Chapron, E. A. & Triplett, R. F. *Proc. Soc. exp. Biol. Med.* **122**, 1167–1171 (1966).
- Davies, A. J. S., Leuchars, E., Wallis, V., Marchant, R. & Elliott, E. V. *Transplantation* **5**, 222–230 (1967).
- Mitchell, G. F. & Miller, J. F. A. P. *J. exp. Med.* **128**, 821–837 (1968).
- Rajewsky, K., Schirmacher, V., Nase, S. & Jerne, N. K. *J. exp. Med.* **129**, 1131–1143 (1969).
- Mitchison, N. A. *Eur. J. Immun.* **1**, 18–27 (1971).
- Katz, D. H., Hamaoka, T. & Benacerraf, B. *J. exp. Med.* **137**, 1405–1420 (1973).
- Schimpl, A. & Wecker, E. *Nature* **226**, 658–660 (1970).
- Dulton, R. W. *et al. Prog. Immun.* **1**, 355–368 (1971).
- Tite, J. P., Kaye, J. & Jones, B. *Eur. J. Immun.* **14**, 553–561 (1984).
- Tite, J. P. & Janeway, C. A. *Eur. J. Immun.* **14**, 878–886 (1984).
- Tite, J. P., Powell, M. B. & Ruddle, N. H. *J. Immun.* **135**, 25–33 (1985).
- Jones, B., Tite, J. P. & Janeway, C. A. *J. Immun.* **136**, 348–356 (1986).
- Janeway, C. A. *et al. J. Immun.* **135**, 739s–742s (1985).
- Janeway, C. A. *et al. in Symposium on Fundamental Cancer Research*, Vol. 38, 45–61 (University of Texas Press, Austin, 1986).
- Jones, B. & Janeway, C. A. *Nature* **292**, 547–549 (1981).
- Tite, J. P. & Janeway, C. A. *J. molec. cell. Immun.* **1**, 253–265 (1984).
- McGee, H. J., Long, S. R. & Briggs, W. R. *Nature* **308**, 667–668 (1984).
- Kaye, J. *et al. J. exp. Med.* **158**, 836–856 (1983).
- Kaye, J. & Janeway, C. A. *J. exp. Med.* **159**, 1397–1412 (1984).
- Kaye, J. *et al. J. Immun.* **133**, 1339–1345 (1984).
- Tite, J. P. *et al. J. exp. Med.* **163**, 189–202 (1986).
- Kupfer, A., Swain, S. L., Janeway, C. A. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6080–6083 (1986).
- Kupfer, A., Singer, S. J., Janeway, C. A. Jr & Swain, S. L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5888–5892 (1987).
- Saizawa, K., Rojo, J. & Janeway, C. A. Jr *Nature* **328**, 260–263 (1987).
- Lichtman, A. H., Kurt-Jones, E. A. & Abbas, A. K. *Proc. natn. Acad. Sci. U.S.A.* **84**, 824–827 (1987).
- Killar, L., MacDonald, G., West, J., Woods, A. & Bottomly, K. *J. Immun.* **138**, 1674–1679 (1987).
- Kim, J., Woods, A., Becker-Dunn, E. & Bottomly, K. *J. exp. Med.* **162**, 188 (1985).
- Rock, K. L., Benacerraf, B. & Abbas, A. K. *J. exp. Med.* **160**, 1102–1113 (1984).
- Lanzavecchia, A. *Nature* **314**, 537–539 (1985).